

**UNIVERSIDAD COMPLUTENSE DE MADRID**

**FACULTAD DE VETERINARIA**



**SECUENCIACIÓN Y ANÁLISIS DEL GEN DE LA  
MIOSTATINA BOVINA**

**MEMORIA PARA OPTAR AL GRADO DE DOCTOR**

**PRESENTADA POR**

Luis José Royo Martín

**Madrid, 2003**

**ISBN: 84-669-2159-1**

**UNIVERSIDAD COMPLUTENSE DE MADRID**



**Secuenciación y análisis del gen de la  
miostatina bovina.**



**TESIS DOCTORAL  
Facultad de Veterinaria  
Madrid, 2002**

**Luis José Royo Martín**

Este trabajo de tesis se ha realizado gracias a la financiación del  
proyecto CICyT AGF98-1087.

# ÍNDICE

---

Índice general.....	i
Índice de Figuras y Tablas.....	vii
Lista de siglas y abreviaturas.....	xi
Resumen/Summary.....	xiii

## Introducción.

1	La hipertrofia muscular en la raza Asturiana de los Valles.....	1
2	Localización del gen responsable de la hipertrofia muscular bovina.....	6
	2.1 Análisis de segregación.....	9
	2.2 Análisis de ligamiento.....	10
	2.3 Análisis de asociación.....	12
3	Clonado posicional del gen responsable de la hipertrofia muscular bovina.....	15
	3.1 Cartografía física.....	17
	3.1.1 Cartografía física de baja resolución.....	18
	3.1.1.1 Paneles de células híbridas somáticas.....	18
	3.1.1.2 Paneles de híbridos de irradiación.....	18
	3.1.1.3 FISH ( <i>Fluorescence in situ hybridization</i> ) ( <i>Chromosome painting</i> ).....	19
	3.1.1.4 Citometría de flujo.....	20
	3.1.2 Cartografía física de alta resolución.....	20
	3.1.2.1 Paseo cromosómico ( <i>Chromosome Walking</i> ).....	21
	3.1.2.2 Cartografía de STS.....	21



3.1.2.3	Uso de ADN repetitivo.....	22
3.1.3	Cartografía física comparada de la región candidata.....	22
3.1.3.1	Cartografía física comparada de la región 2q11-2q22.....	23
3.1.3.2	Construcción de un <i>contig</i> de YAC.....	24
3.2	Cartografía genética.....	25
3.3	Identificación de genes candidatos.....	25
3.3.1	El gen GDF8.....	26
3.3.2	La miostatina como gen candidato responsable del fenotipo culón en bovino.....	27
3.4	Análisis de mutaciones en el gen candidato.....	27
3.4.1	Una mutación en el gen de la miostatina es la responsable de la aparición del fenotipo culón en la raza Blanco-Azul Belga y Asturiana de los Valles	28
3.4.1.1	Características básicas de la miostatina bovina.....	29
3.4.2	El fenotipo culón está causado por una serie alélica de mutaciones del gen de la miostatina bovina.....	30
3.5	Desarrollo de un test diagnóstico de las mutaciones. Selección asistida por marcadores.....	33
	<b>Objetivos.....</b>	<b>35</b>

## Capítulo 1

### Contribución al clonado posicional del gen culón. Reducción del intervalo cromosómico.

1	Elección de YAC para aislar microsatélites.....	37
1.1	Descripción de la genoteca bovina de ADN genómico en YAC.....	37
1.2	Organización y análisis de una genoteca de YAC.....	37
2	Protocolo general de aislamiento de microsatélites.....	39
2.1	Aislamiento del YAC.....	39
2.1.1	Electroforesis en campo pulsado.....	39
2.1.2	Transferencia de ADN a una membrana de nailon.....	41
2.1.3	Hibridación.....	42

2.2	Identificación de microsatélites.....	44
2.2.1	Fragmentación del YAC.....	44
2.2.2	Selección de clones positivos.....	45
2.3	Secuenciación de clones positivos.....	46
3	Resultados.....	50

## Capítulo 2

### Secuenciación del gen de la miostatina bovina y regiones flanqueantes.

1	Antecedentes.....	51
1.1	Tamaño de la secuencia.....	52
1.2	Estrategia de secuenciación.....	52
2	Secuenciación del gen de la miostatina bovina y regiones flanqueantes.....	53
2.1	Obtención de un clon de fago $\lambda$ de la miostatina bovina.....	54
2.1.1	Genotecas bovinas de ADN genómico en fago $\lambda$ .....	54
2.1.2	Siembra de una genoteca de fagos.....	55
2.1.2.1	Titulación de la genoteca.....	56
2.1.2.2	Siembra de la genoteca.....	57
2.1.2.3	Conservación de la genoteca.....	58
2.1.3	Réplicas de la genoteca de fagos.....	58
2.1.4	Aislamiento de clones positivos.....	58
2.2	Clones positivos de la miostatina bovina.....	60
2.3	Amplificación de los clones elegidos.....	61
2.4	Digestión de los productos PCR y clonado de los fragmentos de digestión.....	62
2.5	Secuenciación a partir de los productos PCR.....	65
2.6	Ensamblado de todas las secuencias obtenidas.....	67
3	Estudio de la secuencia.....	69
3.1	Extracción de ARN muscular esquelético.....	69
3.2	Estudio del ARN mensajero.....	69
3.2.1	La secuencia codificante de la miostatina está constituida por tres exones interrumpidos por dos intrones.....	69
3.2.2	Límites de los transcritos.....	69

3.2.2.1 Inicio de la transcripción.....	69
3.2.2.1.1 C-Race.....	70
3.2.2.1.2 Primer extension.....	72
3.2.2.2 Final de la transcripción.....	74

## Capítulo 3

### Identificación de las posibles regiones reguladoras del gen de la miostatina.

1 Antecedentes.....	77
1.1 Identificación de regiones reguladoras de la transcripción.....	77
2 Comparación de las secuencias de la miostatina bovina y de ratón.....	78
3 Búsqueda de regiones reguladoras músculo-específicas.....	81

### Discusión.

1 Reducción del intervalo cromosómico.....	83
2 Estructura del gen de la miostatina bovina.....	84
2.1 Transcritos del gen de la miostatina bovina.....	84
2.2 Estudio del promotor.....	86
3 Estudio de la proteína.....	87
3.1 Estructura molecular de la miostatina.....	88
3.2 Mecanismo de acción biológica.....	90
4 Estudio de las secuencias repetidas.....	91
5 Comparación de las secuencias bovina y murina del gen de la miostatina.....	93
6 Regulación de la transcripción del gen de la miostatina.....	93

Conclusiones.....	95
-------------------	----

Referencias Bibliográficas.....	97
---------------------------------	----

Anexo 1. Soluciones utilizadas.....	107
-------------------------------------	-----

Anexo 2. Marcadores de tamaño de ADN.....	115
---	-----

Anexo 3. Secuencias de la miostatina bovina.....	117
--	-----

#### Anexo 4: Publicaciones más relevantes..... 119

- L. Grobet, **L.J. Royo**, D. Poncelet, D. Pirottin, B. Brouwers, J. Riquet, A. Schoberlein, S. Dunner, F. Menissier, J. Massabanda, R. Fries, R. Hanset, M. Georges (1997). A deletion in the bovine myostatin gene caused the double-muscled phenotype in cattle. *Nature Genetics*, **17**: 71-74.

- L. Grobet, D. Poncelet, **L.J. Royo**, B. Brouwers, D. Pirottin, C. Michaux, F. Menissier, M. Zanotti, S. Dunner, M. Georges (1998). Molecular definition of an allelic series of mutations disrupting the myostatin function and causing double muscled in cattle. *Mammalian Genome*, **9**: 210-213.

- D. Pirottin, D. Poncelet, L. Grobet, **L.J. Royo**, B. Brouwers, J. Masabanda, H. Takeda, R. Fries, K. Sugimoto, J.E. Womack, S. Dunner, & M. Georges (1999). High resolution human-bovine comparative mapping based on a closed YAC contig spanning the bovine mh locus. *Mammalian Genome*, **10**, 289-293.

## Índice de figuras y tablas.

---

### A) Figuras:

- Figura 1. Foto de portada. Lote de 23 culones embarcados para Chile el día 22/2/57. Archivo fotográfico del CENSYRA-Somío..... Portada
- Figura 2. Ternero mestizo culón en Asturias a finales de los años 50. Archivo fotográfico del CENSYRA-Somío..... 2
- Figura 3. Sementales de la estación de Somío. A la izquierda Bolero20 y a la derecha Peter. Archivo fotográfico del CENSYRA-Somío..... 5
- Figura 4. Etapas del clonado posicional..... 8
- Figura 5. Posiciones relativas de los tres genotipos respecto al peso muscular (Hanset y Michaux, 1985a)..... 9
- Figura 6. Localización *locus mh*. Curva de *lodscore* multipunto, análisis de ligamiento (Charlier y col., 1995)..... 11
- Figura 7. Localización *locus mh*. Curva de *lodscore* multipunto, análisis de ligamiento (Dunner y col., 1997)..... 14
- Figura 8. Localización *locus mh*. Curva multipoint, desequilibrio de ligamiento (Dunner y col., 1997)..... 14
- Figura 9. Distribución de la frecuencia alélica del marcador TGLA44 (Dunner y col., 1997)..... 15
- Figura 10. FISH *painting* con sondas humanas y porcinas (Milan y col., 1996)..... 19
- Figura 11. Identificación de cromosomas de cerdo por Citometría de Flujo. (Riquet y col., 1996)..... 20
- Figura 12. Cartografía de la región BTA2q11-2q21 (L. Grobet)..... 25
- Figura 13. Secuencia de ADNc. obtenida a partir de un individuo culón y uno normal, mostrando la delección nt821del(11) (Grobet y col., 1997)..... 28
- Figura 14. Características generales del gen y la proteína la miostatina bovina..... 29
- Figura 15. Representación esquemática del gen de la miostatina bovina con la posición de los polimorfismos encontrados (Grobet y col., 1998)..... 32
- Figura 16. Organización genoteca de YAC (Strachan y Read, 1999)..... 38

- Figura 17. Detalle de los YAC elegidos (en azul) para aislar microsatélites.....	39
- Figura 18. Electroforesis en campo pulsado.....	41
- Figura 19. Actividad de la sonda dependiendo del volumen de elución.....	43
- Figura 20. Gel de agarosa con las PCR de algunos de los microsatélites.....	49
- Figura 21. Amplificación del marcador Bulge05.....	49
- Figura 22. Cartografía de la región con los nuevos marcadores.....	50
- Figura 23. Localización de los oligonucleótidos utilizados sobre la secuencia de la miostatina bovina.....	60
- Figura 24. Características de los clones positivos encontrados.....	60
- Figura 25. Amplificación de los fragmentos de PCR elegidos, en condiciones de “Long Range PCR”.....	61
- Figura 26. Digestión <i>MboI</i> de los productos PCR.....	62
- Figura 27. ADN de cada una de las bandas del patrón de la digestión.....	62
- Figura 28. Localización de los fragmentos obtenidos a partir de secuencias de los clones.....	65
- Figura 29. Avance de la secuencia tras la primera ronda de secuenciación a partir de producto PCR.....	66
- Figura 30. Avance de la secuencia del extremo 3’.....	66
- Figura 31. Avance de la secuencia tras varias rondas de secuenciación.....	67
- Figura 32. Características finales de la secuencia ensamblada.....	68
- Figura 33. Gel de agarosa que muestra la presencia de dos posibles lugares de inicio del ARN mensajero de la miostatina bovina y su secuencia.....	71
- Figura 34. Gel de agarosa donde se muestra la presencia de los 4 posibles lugares de terminación del ARN mensajero del gen de la miostatina.....	76
- Figura 35. Comparación de las secuencias bovina (BTMSTN) y de ratón (MMMSTN).....	79
- Figura 36. Alineamiento de las secuencias homólogas de la miostatina bovina (inferior) y de ratón (superior).....	80
- Figura 37. <i>Contig</i> de YAC de la región bovina de la miostatina (Pirottin y col., 1999).....	84
- Figura 38. Comparación entre las secuencias de la miostatina de ratón y bovina en la región circundante al LP1 bovino.....	86
- Figura 39. Alineamiento en la región del promotor, y localización de la TATA-box y CAAT-box.....	87

-	Figura 40. Estructura molecular de la miostatina bovina.	
-	a) Estructura lineal de la proteína.....	89
-	b) Estructura del homodímero de miostatina, supuesta por analogía a los integrantes de la familia TGF- $\beta$ .....	90
-	Figura 41. Posible mecanismo de acción de la miostatina por analogía a los integrantes de la familia TGF- $\beta$ .....	91
-	Figura 42. Secuencias SINE bovinas.....	92
-	Figura 43. Localización de los elementos repetitivos en la secuencia de la miostatina bovina.....	92

## **B) Tablas:**

-	Tabla I. Comparación general de las características de la canal del ganado culón y normal en porcentaje (Boyajeán y col., 1971).....	2
-	Tabla II. Estadísticos de características cuantitativas de crecimiento en vivo y canal, en la raza bovina Asturiana de los Valles, según presencia o no de hipertrofia muscular (Vallejo y col., 1993).....	3
-	Tabla III. Caracteres productivos y reproductivos en animales culones y normales de la raza Asturiana de los Valles (Cañón y col., 1996).....	5
-	Tabla IV. Herencia del carácter culón.....	9
-	Tabla V. STS en la región BTA2q11-21.....	24
-	Tabla VI. Haplotipos de la miostatina bovina en diferentes razas europeas.....	34
-	Tabla VII. YAC elegidos para aislar microsatélites.....	38
-	Tabla VIII. Características de los clones positivos identificados.....	48
-	Tabla IX. Microsatélites identificados.....	49
-	Tabla X. Secuencias de los oligonucleótidos de los microsatélites identificados....	49
-	Tabla XI. Titulación de la genoteca.....	56
-	Tabla XII. Siembra secundaria en placas de Petri redondas grandes.....	59
-	Tabla XIII. Siembra terciaria.....	59
-	Tabla XIV. Clones positivos de miostatina bovina.....	60
-	Tabla XV. Protocolos de PCR para caracterizar clones positivos.....	61
-	Tabla XVI. Microplaca con clones individuales elegidos.....	63
-	Tabla XVII. Clones elegidos para secuenciar.....	63

- Tabla XVIII. Fragmentos de secuencias a partir de clones.....	64
- Tabla XIX. Oligonucleótidos utilizados en el protocolo de <i>C-RACE</i> .....	71
- Tabla XX. Oligonucleótidos utilizados en el protocolo de 3'- <i>RACE</i> .....	76



## Lista de siglas y abreviaturas.

---

- **A:** adenina
- **ADN:** Ácido desoxirribonucleico.
- **ADNc:** Ácido desoxirribonucleico complementario o codificante.
- **ARN:** Ácido ribonucleico
- **ARNt:** Ácido ribonucleico de transferencia.
- **ASEAMO:** Asociación Española de criadores de ganado de vacuno selecto de raza Asturiana de la Montaña.
- **ASEAVA:** Asociación Española de criadores de ganado de vacuno selecto de Asturiana de los Valles.
- **BAC:** *Bacterial Artificial Chromosome*, cromosoma artificial de bacteria.
- **BBB:** *Blanc Bleu Belge*, Blanco azul belga.
- **C:** citosina.
- **CATS:** *Comparative Anchor Tag Site*, etiqueta de secuencias comparadas.
- **cM:** centiMorgan.
- **cR:** centiRay.
- **dATP:** desoxi adenosina trifosfato.
- **dCTP:** desoxi citosina trifosfato.
- **dGTP:** desoxi guanosina trifosfato.
- **dNTP:** desoxi nucleótido trifosfato.
- **dTTP:** desoxi timina trifosfato.
- ***E. coli:*** *Escherichia coli*.
- **EST:** *Expressed Sequence Tag*, etiqueta de una secuencia codificante.
- **FISH:** *Fluorescence In Situ Hybridization*, Hibridación *in situ* con fluorescencia.
- **G:** guanina.
- **IA:** Inseminación Artificial.
- **IBD:** *Identity By Descent*, idéntico por descendencia.
- **IBS:** *Identity By State*, idéntico por estado.
- **Kb:** kilobase,  $10^3$  bases.
- **LINE:** *Long Interspersed Repetitive Element*, elemento repetitivo largo disperso.
- **MAS:** *Marker Assisted Selection*, selección asistida por marcadores.
- **Mb:** Megabase,  $10^6$  bases.
- **MgCl<sub>2</sub>:** Cloruro de magnesio.
- **mh:** *muscular hypertrophy*, hipertrofia muscular.
- **OLA:** *Oligonucleotide Ligation Assay*, ensayo de ligado con oligos.
- **ORF:** *Open Reading Frame*, marco de lectura abierta.
- **PCR:** *Polymerase Chain Reaction*, reacción en cadena de la polimerasa.
- **Pfu:** plaque formation units, unidad formadora de placas de lisis.
- **QTL:** *Quantitative Trait Loci*, loci responsables de caracteres cuantitativos.
- **RACE:** *Rapid Amplification of cDNA Ends*, Amplificación rápida de los extremos del ADN codificante.
- **rpm:** revoluciones por minuto
- **SINE:** *Short Interspersed Repetitive Element*, elemento repetitivo corto disperso.
- **SRA:** secuencia de replicación autónoma.
- **STS:** *Sequence-Tagged Site*, etiqueta.

- **T:** timina.
- **UCM:** Universidad Complutense de Madrid
- **UP:** Ultra Puro
- **YAC:** *Yeast Artificial Chromosome*, cromosoma artificial de levadura.

## Resumen

---

En diversas especies de mamíferos aparecen individuos con gran desarrollo muscular que en los bovinos se denominan comúnmente culones. Este tipo de animales se encuentran en varias razas y tienen el aprecio de una parte importante de ganaderos y carniceros. Los animales culones se caracterizan por un incremento del orden de un 20% en la masa muscular, debido a una hiperplasia muscular generalizada, es decir un incremento en el número de fibras musculares más que un incremento en el diámetro individual de estas fibras. Además el fenotipo incluye una reducción del 50% del contenido lipídico total así como una reducción en el contenido de tejido conectivo, dando lugar a una carne más tierna y magra muy apreciada por los consumidores. Estas características han provocado el amplio uso de los animales culones para la producción de carne en algunas poblaciones, como la Blanco-Azul Belga en la zona centroeuropea y la Asturiana de los Valles en España. Sin embargo este fenotipo tiene la desventaja de estar asociado a una alta incidencia de problemas en el parto, lo que ha restringido su uso en la mayoría de las poblaciones bovinas.

La hipertrofia muscular es un carácter hereditario sin embargo su modelo de herencia fue causa de controversia, hasta que un estudio de segregación en la raza Blanco-Azul Belga, realizado tanto en cruzamientos experimentales como en poblaciones reales señaló, un modelo de herencia autosomal recesivo. Esta hipótesis se confirmó más tarde cuando se localizó el *locus mh* por análisis de ligamiento en la región centromérica del cromosoma 2 bovino. Posteriormente se identificó el mismo *locus mh* como responsable del fenotipo culón en al menos otras dos razas: la Asturiana de los Valles y la Maine-Anjou.

En el transcurso de este trabajo se ha usado una estrategia de clonado de candidatos posicionales para demostrar, que varias mutaciones en la secuencia del gen de la miostatina son las responsables del fenotipo culón en bovinos. Se han identificado 5 mutaciones causantes del síndrome de la cularidad en varias razas bovinas europeas.

Se ha obtenido una secuencia de ADN de 17.417 bases que incluye la secuencia del ADN codificante, más aproximadamente 5 kb en cada uno de los extremos 5' y 3' del gen. El gen de la miostatina consta de tres exones interrumpidos por dos intrones. Se han identificado en el transcrito de la miostatina los posibles lugares de inicio y final de la transcripción y además se ha caracterizado el promotor del gen.

Utilizando una estrategia de secuenciación comparativa se han identificado 3 regiones en la secuencia del gen de la miostatina, como posibles regiones importantes en la regulación de la transcripción del gen.

## Summary

---

Animals, which exhibit an exceptional muscular development are referred to as double-muscled animals. They have been described in several cattle breeds, and they have received considerable attention from beef producers. Double-muscled animals are characterised by an increase in muscle mass of the order of 20%, which was shown to be due to a generalised skeletal muscle hyperplasia, i.e. an increase in the number of muscled fibers rather than in their individual diameter. In addition a 50% reduction in total lipid content as well as a reduction in connective tissue fraction are shown in animals carrying this phenotype, resulting in leaner and tenderer meat greatly appreciated by the consumers. All these advantages have led to the widespread use of these animals in beef production in some populations such as the Belgian-Blue cattle breed in the Centre of Europe and Asturiana de los Valles cattle breed in Spain. However, the association of calving difficulties with the double-muscled phenotype, has restricted its exploitation in most populations.

Despite the hereditary nature of the trait was assumed early, the precise mode of inheritance has remained controversial until the segregation analysis carried out in Belgian-Blue cattle breed. This study performed both in experimental crosses and in the outbred population pointed towards an autosomal recessive inheritance. This hypothesis was confirmed when the *mh* locus was unambiguously mapped by linkage analysis to the centromeric end of bovine chromosome 2. Later, the same *mh* locus was shown to underly double-muscling in at least two others breeds: Asturiana de los Valles and Maine-Anjou.

During our work, a positional candidate cloning approach has shown that several mutations in the myostatin gene are responsible for the doubled-muscled phenotype, and up to 5 different mutations have been described in different european breeds.

A 17.417 pb long DNA sequence of the myostatin gene has been obtained, which includes all the complete gene, plus 5 kb in both 5' and 3' extremes. Myostatin gene is known to have three exons interrupted by two introns. The putative initial and final points of transcription have been identified, and the myostatin promoter has been characterised as well.

By means of a comparative sequencing approach, three conservative DNA regions have been identified as important in the transcription regulation of the myostatin gene.

## Introducción

---

### 1- La hipertrofia muscular en la raza Asturiana de los Valles.

La producción de carne de bovino en el continente europeo durante el siglo XX mediante razas especializadas, se ha visto condicionada por la necesidad de los ganaderos de satisfacer una demanda de los consumidores de una carne tierna, clara y magra, y de obtener animales con alto potencial de crecimiento muscular. Esta situación ha propiciado una selección a favor de aquellos animales que presentaban un fenotipo concreto, caracterizado por un gran desarrollo muscular generalizado. Este fenotipo ha recibido diferentes nombres: grupa doble, culones, *cul de poulain*, *croupe de poulain*, *culard*, *paardenbil*, *dikbil*, *doppellender*, *double-muscled*, *groppa di cavallo*, *groppa dopia*.

La hipertrofia muscular hereditaria es una característica genética mencionada por primera vez por Culley en 1807, en Inglaterra en ganado Shorthorn. A principios de siglo se describe en el norte de Alemania y más tarde en varias razas europeas, incluida la Frisona. Actualmente, el carácter culón tiene difusión mundial, y se presenta en numerosas razas como la Charolaise, Maine d'Anjou, Blonde d'Aquitaine, Parthenaise, Tarantaise, Bazadaise y Limousin en Francia; Piamontese en Italia; Blanc-Bleu en Bélgica; South-Devon en Gran Bretaña; Sta. Gertrudis, Angus Americano, Galloway y Hereford en EE.UU.; Timina en Cuba; Sta. Gertrudis en Australia y finalmente en España en la Rubia Gallega, Asturiana de los Valles y Pirenaica.

Solo cuando se comparan tipos extremos y bien diferenciados de animales con y sin hipertrofia muscular (culones-normales), se pueden discernir claramente las características que diferencian externamente a los animales culones. Los animales culones presentan un aspecto completamente diferenciado respecto de los normales, debido especialmente a una hipertrofia de las masas musculares en gradiente positivo antero-posterior y disto-proximal (Ménissier, 1982), especialmente en las áreas crurales y braquiales (Boccard y Dumont, 1974); la masa muscular se mantiene e incluso se observa hipotrofia en las regiones costo-torácica, abdominal y del cuello, presentando un aspecto más compacto. Las masas musculares y los surcos intermusculares son especialmente visibles bajo la piel, lo que sugiere una fuerte reducción de la capacidad de depósito de grasa subcutánea, probablemente unida a una alteración del metabolismo lipídico (Holmes y Robinson, 1970).

La hipertrofia muscular se asocia a un menor desarrollo óseo, que en la región raquídea parece seguir un gradiente negativo antero-posterior (McKellar, 1968), y a un menor desarrollo del aparato digestivo (Boyajeau y col., 1971). El aparato digestivo de los animales culones puede tener un peso 13% menor que el de los animales normales, provocando un “agalgamiento del vientre”, lo que en conjunto da a los animales culones cierto aspecto de caballo (grupa de potro) (Ménissier, 1982).

Estas diferencias externas *in vivo* se reflejan también en la canal, siguiendo los mismos gradientes. Respecto de las canales de los animales normales, los culones presentan menos grasa intramuscular (West, 1976a) y un color más pálido (Boccard y Monin, 1973). Boyajeau y col., (1971) revisan las características de las canales que presentan hipertrofia muscular hereditaria respecto de las normales (Tabla I), mostrando un mayor porcentaje de rendimiento a la canal (+8% en valor relativo), una amplia reducción del hígado, corazón y pulmón (hasta un -12%), de la piel y el aparato digestivo (-13%), y una mayor proporción de músculo en la canal (+17% en valor relativo), especialmente si se compara con las reducciones en hueso y grasa de la canal (-14 % y -44 % en valor relativo respectivamente). La calidad de la canal también se ve afectada, ya que la carne de los animales culones tiene un mayor contenido en agua (Raimondi, 1957), peor textura (Kidwell y col., 1952) y mayor terneza (West, 1976b) que la de los animales normales.

- Tabla I. Comparación general de las características de la canal del ganado culón y normal en porcentaje (Boyajeau y col., 1971).

Carácter	Tipo de Animal y ponderación			
	C <sup>1</sup>	N <sup>1</sup>	C-N <sup>1</sup>	(C-N)*100/N <sup>2</sup>
Cabeza	3,5	3,6	-0,2	-2,0
Patas	2,6	2,8	-0,2	-6,3
Piel	6,1	7,3	-1,2	-13,0
Corazón	0,4	0,4	0,0	-12,0
Pulmón	0,8	1,0	-0,1	-12,1
Hígado	1,2	1,3	-0,2	-12,2
Aparato digestivo	4,5	5,5	-1,0	-13,1
Rendimiento canal	64,4	59,9	4,5	8,2
% Músculo canal	75,3	68,3	11,5	17,4
% Grasa canal	7,9	17,7	-9,8	-43,5
% Hueso canal	13,4	15,8	-2,4	-13,9
Relación músculo hueso	5,87	4,17	1,70	35,3

C: Culón; N: Normal. 1: Media no ponderada de datos bibliográficos. 2: Media experimental ponderada como  $2xN_c \times N_n / (N_c + N_n)$ :  $N_c$  Número de animales culones;  $N_n$ : Número de animales normales.

En la raza Asturiana de los Valles, los animales con fenotipo culón, a edades similares a los normales, presentan mejores canales y mejores características cuantitativas y de composición, aunque no presentan ninguna superioridad en relación con su crecimiento. Las canales de animales culones son más cortas y menos profundas, pero al ser más compactas y con mayor proporción de músculo, su conformación es mejor y se califican en categorías superiores (Tabla II; Vallejo y col., 1993).

Histológicamente el mayor desarrollo de los músculos de un animal culón es debido sobre todo a un incremento en el número de fibras musculares (hiperplasia), más que un aumento individual de las fibras musculares (Hanset y Michaux, 1982). Los animales culones poseen al nacimiento casi el doble de fibras musculares que los animales normales (Gerrard y col., 1991). También se ha encontrado que los animales culones tienen un contenido menor en tejido conectivo (colágeno) (Boccard 1982; Uytterhaegen y col., 1994), y esta reducción se estima del 20-30 %, utilizando los

niveles de hidroxiprolina como indicadores del contenido de colágeno en el músculo (Hanset y Michaux, 1982).

- Tabla II. Estadísticos de características cuantitativas de crecimiento en vivo y canal, en la raza bovina Asturiana de los Valles, según presencia o no de hipertrofia muscular (Vallejo y col., 1993).

Variables	Hipertrofia muscular (n :6)		Tipo normal (n:21)		
crecimiento y canal	m	DT	m	DT	F
Edad(meses)	14,50 ± 0,55		14,91 ± 1,10		0,76
PV(kg)	494,83 ± 67,37		520,41 ± 63,19		0,75
GMD(kg)	1,12 ± 0,17		1,10 ± 0,10		0,07
ICV	4,44 ± 1,32		4,39 ± 0,84		0,01
PC(kg)	313,67 ± 42,46		300,55 ± 42,12		0,46
R(%)	63,41 ± 0,45		57,64 ± 3,04		20,90***
LC(cm)	124,67 ± 10,07		128,52 ± 5,73		1,49
LP(cm)	79,83 ± 4,26		83,86 ± 8,01		1,37
PP(cm)	56,17 ± 7,41		57,86 ± 2,82		0,77
ICC	2,52 ± 0,22		2,34 ± 0,27		2,24
MC(kg)	127,42 ± 18,82		110,88 ± 17,41		4,07
HC(kg)	22,15 ± 2,96		25,80 ± 3,29		5,99*
GC(%)	4,10 ± 0,99		9,85 ± 3,97		12,06***
PMC(%)	82,73 ± 1,53		75,51 ± 3,06		30,58***
PHC(%)	14,44 ± 1,52		17,71 ± 1,82		16,16***
PGC(%)	2,71 ± 0,85		6,66 ± 2,40		15,23***
RMHC	5,78 ± 0,74		4,31 ± 0,56		27,92***
GP	0,37 ± 0,11		1,01 ± 0,50		9,52***
SLD(cm <sup>2</sup> )	154,92 ± 26,43		103,35 ± 28,56		12,63***
m: Media aritmética; DT: Desviación típica; * p< 0,05; *** p< 0,005					

La característica más negativa de la cularidad es su influencia sobre los caracteres reproductivos y maternos de las vacas de cría. Los animales culones presentan mayor dificultad al parto, descenso de la producción de leche y menor precocidad sexual (Ménissier, 1982). Además, este síndrome se asocia con una serie de problemas como reducción de la fertilidad, distocias, baja viabilidad de los terneros y un incremento en la susceptibilidad al estrés (Arthur, 1995).

En la expresión del síndrome de la hipertrofia muscular hay una gran variabilidad, por lo cual a veces se hace difícil la identificación de los animales como culones o normales. Esta expresión también varía durante la vida del mismo animal. La identificación de estos animales normalmente se ha hecho sobre la base de una medida subjetiva del grado de hipertrofia muscular y de otras características externas asociadas al síndrome. También se han utilizado registros genealógicos, cuando se disponía de ellos; e incluso algunas medidas semi-objetivas, como por ejemplo ciertos índices acumulativos de valores basados en la expresión de ciertas características, o la combinación de algunas características de la canal y ciertos parámetros bioquímicos. Pero todos estos métodos sólo sirven para clasificar bien a los animales con fenotipos extremos. Esta variabilidad en la expresión del gen y la falta de un método objetivo de

identificar los animales portadores del gen, ha contribuido a que haya sido difícil descifrar correctamente el modo de acción de este gen.

Se buscó algún método biológico para identificar a los animales portadores del gen de la hipertrofia muscular, como existe por ejemplo en el síndrome de estrés porcino, donde los individuos portadores del gen son sensibles al Halotano (Rempel y col., 1993); sin embargo, no se encontró ningún método que diferenciara a los animales correctamente. Algunos de estos métodos se basaban en las diferencias, entre animales culones y normales, en el tipo de fibras musculares (Holmes y Ashmore, 1972), inervación motora de los músculos (Swatland, 1973; Novakofski y col., 1981), concentraciones de hormona tiroidea (Novakofski y Kauffman, 1981), y concentraciones de creatina y creatinina en la sangre (Hanset y Michaux, 1982; Masoero, 1982).

El carácter culón se manifiesta en la raza Asturiana de los Valles con una frecuencia alta (Cañón y col., 1996). La expresión del carácter en la población Asturiana de los Valles provoca un cambio sustancial de los sistemas de explotación, del manejo y de la propia gestión económica de la ganadería. Los animales de tipo culón tienden a desplazar a los de tipo tradicional y modifican los modos de actuación del ganadero que elige este tipo de animales.

La introducción de la hipertrofia muscular hereditaria en la raza Asturiana de los Valles no se conoce con exactitud ya que no existen referencias a principio del siglo XX sobre el carácter en la población asturiana (Villa, 1999). Es posible que durante el primer tercio del siglo XX, el carácter culón se hubiese difundido ya por toda la geografía regional, y por ejemplo García Fierro (1972) describe la presencia de un semental culón en el año 1941 en el concejo de Llanes, que presentaba una capa berrenda en negro (Figura 2). El mismo García Fierro (1972) atribuye la aparición del carácter en Asturias al cruzamiento de Rubia Gallega con machos Simmental y al gran proceso de mestizaje sufrido por la cabaña ganadera de la cornisa Cantábrica con ganado lechero Frisón en la década de los cuarenta. El ganado Sorthorn, que parece haber tenido un papel esencial en la difusión de la cularidad en el continente europeo (Ménissier, 1982), no ha tenido una presencia importante en Asturias, aunque existen referencias de la importación de animales de esta raza a principios de siglo (Naredo y Bajo, 1916). La creciente mecanización del campo asturiano redujo la necesidad de utilizar animales para el trabajo (donde un animal culón estaba peor dotado), pasando entonces a cobrar mayor importancia la producción de carne en la economía de las explotaciones.

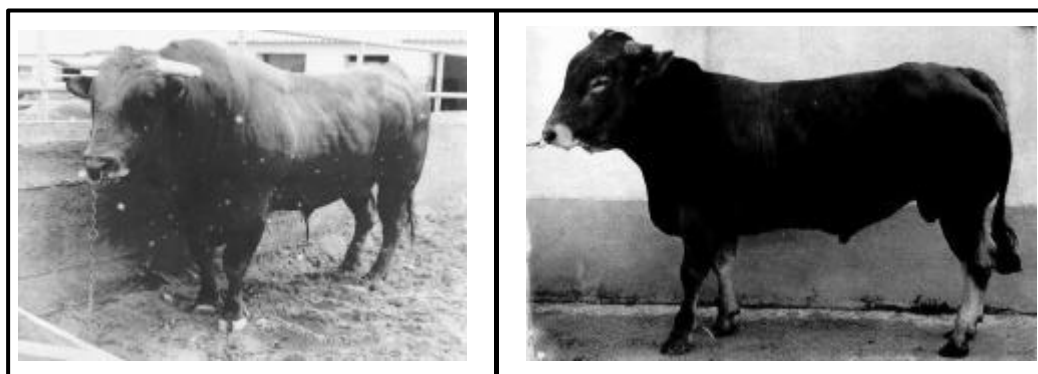
- Figura 2. Ternero mestizo culón en Asturias a finales de los años 50. Archivo fotográfico del CENSYRA-Somío.





Otros factores también debieron contribuir a su difusión. A partir de los años 50 se empezó a emplear la Inseminación Artificial en Asturias, y gracias a la influencia ejercida por el sector comercial, el ganadero demandaba sementales portadores del carácter culón. La estación pecuaria de Somió adquirió para inseminación dos sementales, que presentaban el carácter culón (“Bolero” y “Peter”), a partir de los cuales se crearon estirpes de reproductores que transmitían la hipertrofia muscular.

- Figura 3. Sementales de la estación de Somió. A la izquierda Bolero20 y a la derecha Peter. Archivo fotográfico del CENSYRA-Somió.



Por otra parte, los ganaderos animados por el valor económico superior de estos animales utilizaban, tanto en pureza como en cruzamiento industrial sobre el vacuno lechero, sementales culones para aumentar la rentabilidad de las explotaciones. El último eslabón dentro de la cadena comercial, tratantes y carniceros, apreciaban de una forma especial este tipo de animales ya que, a parte de su mayor peso, proporcionaban un elevado rendimiento útil y comercial en cantidad de masa muscular y piezas nobles. Esto supuso un fuerte incentivo para los ganaderos que reponían en muchos casos con hembras culonas, aumentando la presencia del carácter en la raza Asturiana de los Valles.

- Tabla III. Caracteres productivos y reproductivos en animales culones y normales de la raza Asturiana de los Valles (Cañón y col., 1996).

	Culones	normales
Peso Nacimiento (Kg.)	43,5	39,7
Peso Destete <sup>180</sup> (Kg.)	218	221
Peso Sacrificio <sup>417</sup> (Kg.)	457	509
Rendimiento Canal	63%	58%
% músculo	83%	76%
% hueso	14,4%	17,7%
% grasa	2,7%	6,7%
Superficie <i>Long. dorsi</i>	155cm <sup>2</sup>	103cm <sup>2</sup>
Dificultad al Parto	10,1% <sup>a</sup>	8,8%
12,1% <sup>b</sup>	2,5%	
Edad 1 <sup>er</sup> parto	969 días	931 días
Intervalo entre partos	380 <sup>a</sup>	373
	385 <sup>b</sup>	375

a= el ternero es el animal culón

b= la vaca es el animal culón

## 2- Localización del gen responsable de la hipertrofia muscular bovina.

El progreso en mejora ganadera durante el último siglo ha sido enorme. Las principales razones por las cuales se han podido intensificar las producciones han sido: la aparición de programas de mejora basados en la genética cuantitativa, la utilización de nuevas técnicas reproductivas (inseminación artificial y transferencia de embriones), y la informatización de las ganaderías. Además la revolución tecnológica, incluyendo la biotecnología y la genética molecular, permitirá muchos más avances. La unión de la genética cuantitativa y molecular, permite establecer nuevas estrategias para poder identificar *loci* responsables de caracteres importantes en ganado vacuno, los conocidos como QTL (*Quantitative trait loci*). Un QTL es una región o fragmento de un cromosoma que explica un porcentaje de la varianza aditiva de un carácter productivo. Un caso extremo de un QTL, es cuando ese fragmento de cromosoma explica gran parte o la mayoría de la varianza de un carácter, en ese caso se puede hablar de un gen mayor. Actualmente, con el progreso de la genética molecular, es posible pensar en estrategias que permitan buscar e identificar mutaciones en la secuencia nucleotídica de genes candidatos, responsables de los diferentes fenotipos para un determinado carácter.

Para llegar al clonado posicional de un gen cualquiera hay que seguir una serie de etapas sucesivas (revisadas en Charlier, 1999), que se resumen en la Figura 4. El esquema que se va a describir no tiene porqué ser seguido al pie de la letra y alguna de las etapas puede ser pasada por alto:

1) Definición del fenotipo del carácter: los criterios con que se van a fenotipar los animales, a que edad, si se debe repetir el fenotipado, o si se hace o no sobre el animal vivo.

2) Establecimiento del determinismo genético del carácter: si es heredable, si se trata de un carácter monogénico o poligénico, su modo de transmisión, etc.

3) Obtención de familias informativas para la segregación del carácter. Se recolectan las muestras de los individuos para la extracción del ADN (lo más normal son muestras de sangre o esperma). Además utilizando los mapas de marcadores genéticos publicados, se elegirán una serie de marcadores polimórficos repartidos uniformemente por todo el genoma. Los individuos de las familias informativas son genotipados con los marcadores elegidos. Generalmente se deben corregir los datos del genotipado, por ejemplo de los errores de muestreo o de paternidad. Ahora se realiza el estudio de ligamiento propiamente dicho. En el mejor de los casos el estudio de ligamiento debe dar como resultado una primera localización del gen en cuestión.

4) Confirmación de la localización inicial, bien por introducción de nuevos marcadores informativos en la región, o bien por el análisis de más individuos. Si los marcadores de la región son suficientemente polimórficos y están localizados cerca del *locus* de interés, puede llevarse a cabo una selección asistida por marcadores (MAS: *Marker Assisted Selection*), como por ejemplo en la enfermedad de Weaver (Georges y col., 1993a).

5) Definición de un intervalo cromosómico lo más pequeño posible que contenga el *locus* (del orden de pocos centimorgan), de tal forma que permita la construcción de un mapa físico. Para conseguirlo se buscan nuevos marcadores de la

región utilizando diferentes métodos (cartografía comparada, microdissección, clonado substractivo...), y se añaden nuevas meiosis informativas siempre que sea posible. Si no, se pueden utilizar otros métodos alternativos como por ejemplo explotar el desequilibrio de ligamiento (como en el caso de la hipertrofia muscular bovina, Dunner y col., 1997).

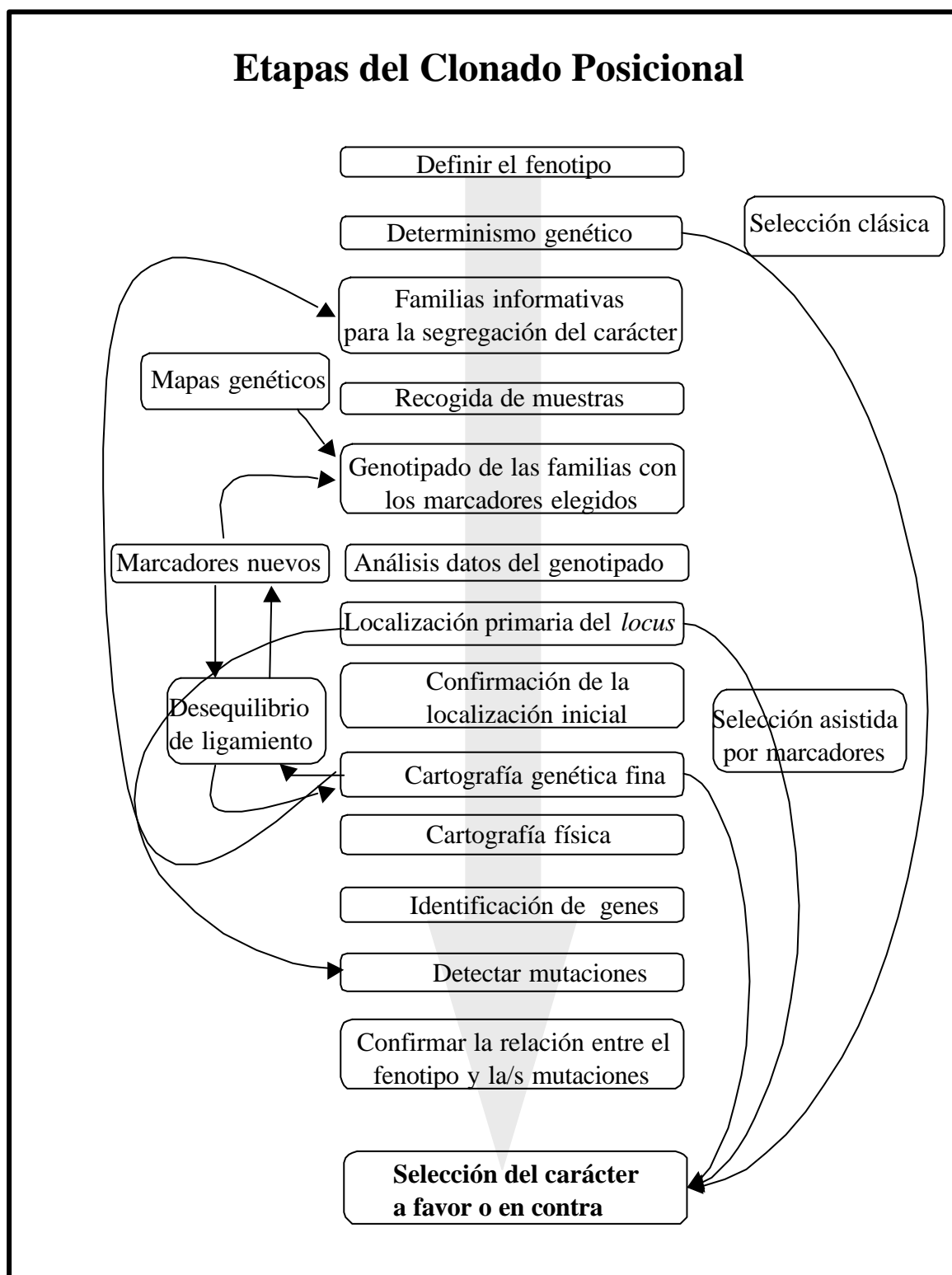
6) Cartografiado físico. La cartografía física puede desarrollarse una vez que se ha definido un intervalo pequeño. Se empieza construyendo un *contig* de clones con insertos de ADN genómico de gran tamaño, utilizando los marcadores o etiquetas (STS, *Sequence-Tagged Site*) de la región para ordenar los clones. También pueden utilizarse otras herramientas como los híbridos de radiación.

7) Identificación de genes dentro del mapa físico. Estos genes, debido a que están localizados dentro del *contig*, serán considerados como candidatos para el carácter buscado. Quizás dentro de este *contig* se encuentren genes de función conocida y con alguna relación con el carácter en estudio, siendo por tanto los candidatos que hay que estudiar en primer lugar.

8) Análisis de mutaciones en los genes identificados dentro del *contig*. Si se detecta una mutación, será necesario demostrar una relación entre el fenotipo del carácter y la mutación identificada, para evitar confundirlo con polimorfismos neutros, que co-segreden con el carácter pero sin relación con él.

9) Desarrollo de un test diagnóstico. Una vez caracterizada la mutación, se desarrolla un test diagnóstico que permita esquemas de selección del carácter estudiado.

- Figura 4. Etapas del clonado posicional.



## 2.1 Análisis de segregación.

Durante la primera mitad del siglo XX, se hicieron siete estudios independientes para intentar descifrar la herencia del carácter culón (Tabla IV), obteniendo una gran variabilidad en los resultados. Sin embargo la mayoría se decantaba por que la causa de la hipertrofia muscular bovina fuera un carácter monogénico.

- Tabla IV. Herencia del carácter culón.

Autores	Razas	Nº Genes	Herencia
Wriedt, 1929	Shorthorn danés	1	Dominancia variable
Magliano, 1933	Piamontese	1	Parcialmente recesiva
Kronacker, 1934	Pie-noir	3	2 para el carácter 1 supresor
Weber e Ibsen, 1934	Hereford U.S.	1	recesiva incompleta, expresión variable, con modificadores
Paci, 1935	Piamontese	1	dominante, expresión variable
Smith, 1949	Aberdeen Angus U.S.	1	recesiva incompleta
Kidwell et al., 1952	Aberdeen Angus x Afrikaander U.S.	1	recesiva incompleta expresión variable.

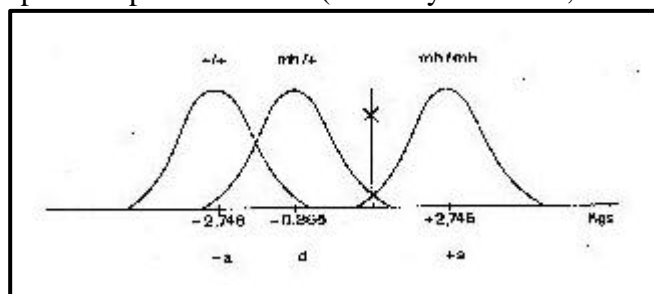
Para responder definitivamente a la pregunta del tipo de herencia Hanset y Michaux (1985 a y b), en la raza Blanco-Azul Belga (BBB, *Blanc-Bleu-Belge*) llevaron a cabo unos análisis de segregación:

Se realizaron cruzamientos dirigidos, con el fin de crear unas hembras F1 de un cruce macho BBB x hembras frisonas, y hacer un retrocruzamiento con toros BBB.

En estos cruzamientos se tenía en cuenta el grado de la musculatura (medido como el peso total de músculos en media canal, después de disecados los músculos de terneros sacrificados a peso constante de 84 kg.). A estos datos, pensando en un determinismo monogénico, se les ajustó un modelo de dos alelos por el método de mínimos cuadrados.

Los resultados obtenidos permitieron concluir (Hanset y Michaux, 1985a) que el carácter culón se ajustaba a un modelo de herencia de un gen mayor parcialmente recesivo, con una serie de genes modificadores, que explicaban la variación dentro de cada grupo de animales. Así, se pudo hablar propiamente de un *locus mh* (*muscular hypertrophy*, aunque en realidad se trate de una hiperplasia), como responsable de este carácter.

- Figura 5. Posiciones relativas de los tres genotipos respecto al peso muscular (Hanset y Michaux, 1985a).



Por otro lado Hanset y Michaux clasificaron hembras reproductoras de blanco-azul belga como “mixtas” o “cárnicas” según su conformación externa, siendo las cárnicas las mejor conformadas. Los descendientes también se clasificaron en estas dos clases en una primera visita a la edad de 1 a 2 meses. Con 11 meses, en la segunda visita, la conformación de los descendientes se evaluó con una nota que va desde 55 hasta 120, de 5 en 5. Las distribuciones de esta nota en la población de descendientes de I.A. (Inseminación Artificial), de los cruces cárnico x cárnico, cárnico x mixto, mixto x mixto, dieron curvas compatibles con la segregación de un gen mayor (Hanset y Michaux, 1985b).

Como resumen a estos trabajos basados en análisis de segregación, se pudo decir que debía existir un *locus* autosómico *mh* caracterizado por un alelo salvaje + y un alelo recesivo *mh* causante en homocigosis del fenotipo culón. Los animales heterocigotos *mh/+* tenían un fenotipo más cercano al homocigoto normal, pero mostraban un cierto grado de hipertrofia muscular, por lo que se asignó al *locus mh* un modelo de herencia parcialmente recesiva. Además, debido a que las diferencias entre las medias de las distribuciones de los animales postulados como *mh/mh* y *+/+* eran de hasta 4 desviaciones típicas en algunas medidas de musculatura, se calificó al gen responsable como un gen de efecto mayor.

Reforzando esta hipótesis se identificó una banda de un patrón multilocus de ADN *fingerprinting* que co-segregaba con el fenotipo culón en la descendencia de un macho heterocigoto (Georges y col., 1990). Sin embargo este posible ligamiento no se encontró en otras familias, y no fue posible ni el clonado de esta banda, ni su localización cromosómica.

## 2.2 Análisis de ligamiento.

Los análisis de ligamiento se basan en estimar el valor de  $\theta$  (fracción de recombinación) entre dos *loci* y determinar si es significativamente distinto de 0,5. El análisis consiste en determinar la probabilidad relativa de obtener los datos observados si los dos *loci* están ligados a una fracción de recombinación determinada  $\theta$ , comparándola con la probabilidad de obtener esos datos si los *loci* no están ligados. En la práctica se calculan una serie de cocientes de probabilidades para distintos posibles valores de  $\theta$ , desde  $\theta=0,00$  (no recombinación) hasta  $\theta=0,50$  (segregación al azar):

Prob. de los datos si los *loci* están ligados a  $\theta < 0,5$  / Prob de los datos si los *loci* no están ligados  $\theta=0,5$

Estas probabilidades se suelen expresar como el  $\log_{10}$  de este cociente y se llama *lod score* = Z (*Logarithm of the Odds* = logaritmo de las probabilidades). Convencionalmente se considera que un *lod score* de 3,0 o mayor constituye una evidencia de ligamiento (equivale a una probabilidad mayor de 1000:1 a favor del ligamiento). El valor de  $\theta$  para el cual Z es máximo se acepta como la mejor estima de la fracción de recombinación.

El análisis de ligamiento puede ser más eficaz cuando más de dos *loci* son analizados simultáneamente. El análisis multilocus es particularmente útil para establecer el orden de una serie de marcadores en un cromosoma. Además tiene una

segunda ventaja: aumentar la información de los marcadores dentro de una familia. Algunas meiosis pueden ser informativas para el marcador A, y otras no informativas para A pero si para el marcador siguiente B. Sólo un análisis de ligamiento simultáneo entre un carácter y los marcadores A y B extrae la información completa.

El análisis de ligamiento ha permitido localizar cromosómicamente varios genes en el ganado bovino, como por ejemplo los genes responsables de: la enfermedad de Weaver (cromosoma 4, Georges y col., 1993a), la ausencia o presencia de cuernos (cromosoma 1, Georges y col., 1993b), la hipertrofia muscular (cromosoma 2, Charlier y col. 1995), el factor rojo (cromosoma 18, Klugland y col, 1995, Charlier 1999), la enfermedad de Heifer (cromosoma 5, Charlier y col., 1996a) y la sindactilia (cromosoma 15, Charlier y col., 1996b).

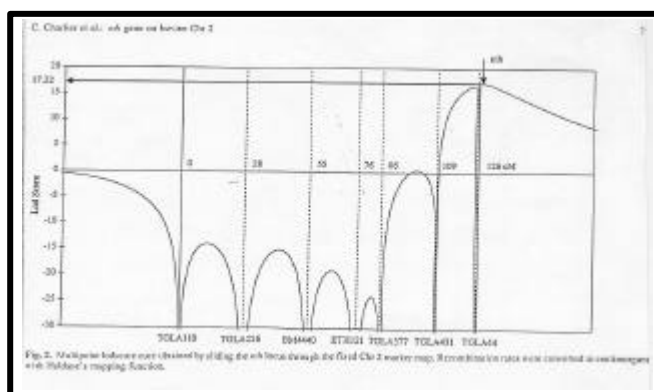
Para localizar el *locus mh* dentro del genoma bovino (mapeo o cartografía genética) se utilizaron marcadores de tipo microsatélite distribuidos a lo largo de los 29 cromosomas autosómicos (exceptuando los cromosomas sexuales).

Se utilizó un esquema basado en un retrocruzamiento con animales de la raza Blanco-Azul Belga (familia de Sart-Tilman). Las familias informativas se generaron cruzando 6 toros hipermusculados (supuestos *mh/mh*) con hembras Holstein (suponiendo un genotipo *+/+*). Posteriormente, se cruzaron 4 de los machos blanco-azul belgas con 41 hembras resultantes del cruce anterior (por lo tanto *mh/+*), produciendo lo que se llama un retrocruzamiento o *bakcross*, de donde se originó una descendencia de 108 terneros. Además para verificar los análisis de ligamiento realizados, se utilizaron 3 familias de medios hermanos paternos (*paternal half-sibs*) obtenidas en explotaciones comerciales.

Estas familias se genotiparon con 213 marcadores tipo microsatélite distribuidos por los 29 autosomas (cromosomas no sexuales). Los resultados más destacables de este estudio fueron los siguientes:

- 1- La descendencia fue de 53 animales hipermusculados y 55 convencionales, lo que indicaba un tipo de segregación mendeliana monofactorial, y de distribución uniforme en ambos sexos.
- 2- Se encontró una fuerte evidencia de ligamiento del *locus mh* con dos marcadores del cromosoma 2 bovino, TGLA431 ( $z = 7,9$ ) y TGLA44 ( $z = 15,8$ ), y posible con otros 3 marcadores más del mismo cromosoma.

- Figura 6. Localización *locus mh*. Curva de *lodscore* multipunto, análisis de ligamiento (Charlier y col., 1995).



Estos datos permitieron afirmar que el *locus mh*, en familias de la raza Blanco-Azul belga se localizaba en el cromosoma 2, a 3 cM. del marcador TGLA44 (Charlier y col., 1995).

### 2.3 Análisis de asociación.

Estos estudios buscan segmentos cromosómicos que son compartidos por los individuos que presentan el fenotipo. Hay que distinguir primero entre segmentos idénticos por descendencia (IBD, *Identity By Descent*), y segmentos idénticos por estado (IBS, *Identity By State*). Dos hermanos pueden tener un genotipo igual para dos marcadores, esto es un segmento idéntico por estado, pero puede o no ser idéntico por descendencia. Puede que uno o los dos padres no tengan ese segmento en el mismo cromosoma, y los hijos hayan heredado un alelo de cada cromosoma, entonces no se trata de un fragmento idéntico por descendencia.

Los segmentos compartidos pueden ser usados dentro de familia, o en poblaciones que descienden de un pequeño grupo fundador. Así, muchas generaciones y por lo tanto muchas meiosis, los separarán de su antecesor común, y durante este tiempo la recombinación puede haber reducido, el segmento compartido, a una pequeña región. En esa región habrá una asociación, a nivel de la población, entre el carácter y un particular alelo o haplotipo

Es decir, se pueden ver los análisis de ligamiento en familias y los estudios de asociación como dos extremos de una misma línea. Cuantas más generaciones separen a dos individuos de su antecesor común, más pequeño será el segmento compartido. Cuando los segmentos son muy pequeños son más difíciles de encontrar, más marcadores se deben usar y el riesgo de falsos positivos es mayor; pero una vez encontrado, podemos localizar el *locus* del carácter con gran precisión. Los estudios de asociación se pueden utilizar tanto para localizar *loci* mendelianos como no mendelianos. Para casos de herencia mendeliana la mejor estrategia, en general, es localizar el *locus* rápidamente con estudios de ligamiento en familias, y entonces usar la asociación entre alelos para afinar la posición. Sin embargo, los caracteres donde las mutaciones son frecuentes no muestran asociación alélica, y sólo pueden ser mapeadas dentro de familias.

Estos métodos se han aplicado, además de en la identificación del *locus mh* (Dunner y col., 1997), para mapear el locus responsable de la sindactilia bovina en el cromosoma 15 (Charlier y col., 1996b). También se han aplicado a la producción de leche, confirmando la localización de un QTL (*Quantitative Trait Loci*, *loci* responsables de caracteres cuantitativos) en la región 14q11-16 bovina. Basándose en técnicas de IBD y mapeo fino comparativo se ha identificado un segmento de 5 cM. que debe contener un QTL. El haplotipo compartido se puede utilizar en estrategias de selección asistida por marcadores (Riquet y col., 1999).

El fenómeno de la hipertrofia muscular estaba descrito en varias razas. La teoría más aceptada del origen de la hipertrofia muscular, era la posible migración de una única mutación a partir del ganado Shorthorn (Ménissier, 1982). Para poder utilizar estos datos, obtenidos en la raza Blanco-Azul belga, con propósitos zootécnicos (selección a favor o en contra del gen, identificación portadores, etc.), era necesario



saber que papel jugaba el *locus mh* en otras razas donde la hipertrofia muscular también estaba descrita. Es decir, continuando con esta teoría, las razas donde se producía la hipertrofia muscular debían compartir un ancestro común, y además un segmento cromosómico idéntico por descendencia en los alrededores del gen (Dunner y col., 1997).

Para ello se utilizaron los marcadores del cromosoma 2 bovino en familias de la raza Asturiana de los Valles. En este trabajo se realizaron dos tipos de análisis: de ligamiento y de desequilibrio de ligamiento.

Para los estudios de ligamiento se eligieron 7 machos de genotipo supuestamente heterocigoto (*mh/+*), deducido de su fenotipo normal y tener descendientes culones (asumiendo un modelo de herencia recesivo como en el Blanco-Azul belga). Se genotipó la descendencia, compuesta por 33 individuos de fenotipo culón y 55 de fenotipo normal, así como las madres si estaban disponibles.

En el caso de los análisis de desequilibrio de ligamiento se muestrearon 27 machos no emparentados hipermusculados y 17 machos no emparentados de fenotipo convencional (basados en la ausencia de individuos fenotípicamente culones en su descendencia).

Los resultados más destacables fueron:

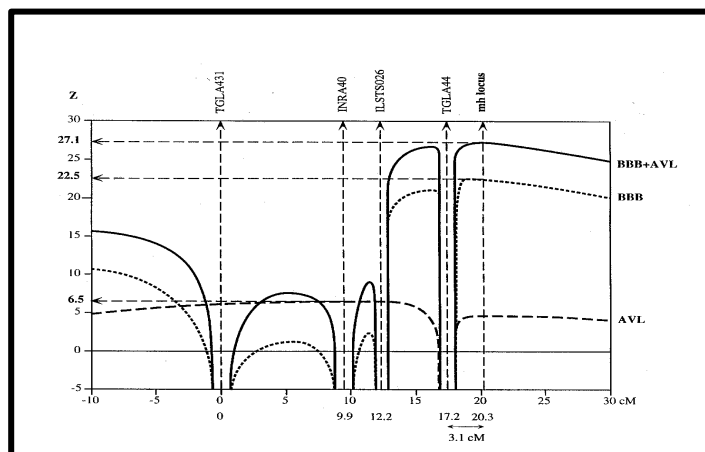
- 1- Se encontraron *lodscore* significativos en las dos razas en las proximidades del marcador TGLA44, aunque en lados opuestos, en Blanco-Azul belga más centromérico y en Asturiana entre TGLA44 e ILSTS026. Sin embargo los intervalos de confianza obtenidos en ambas razas eran superponibles.
- 2- La curva de *lodscore*, utilizando familias de las dos razas juntas, tenía su máximo a 3,1 cM. centromérico con respecto al marcador TGLA44.
- 3- El análisis multipunto de desequilibrio de ligamiento realizado en ambas razas, situaba el *locus* en posiciones muy parecidas al análisis de ligamiento, cercano a TGLA44.
- 4- Se evidenció un desequilibrio de ligamiento entre los marcadores del cromosoma 2 y el *locus mh* en las dos razas. Sin embargo a pesar de haber encontrado un desequilibrio de ligamiento con TGLA44 en las dos razas, el alelo del marcador preferentemente asociado con el cromosoma *mh* era diferente (Dunner y col., 1997).

Las conclusiones de estos estudios de asociación son:

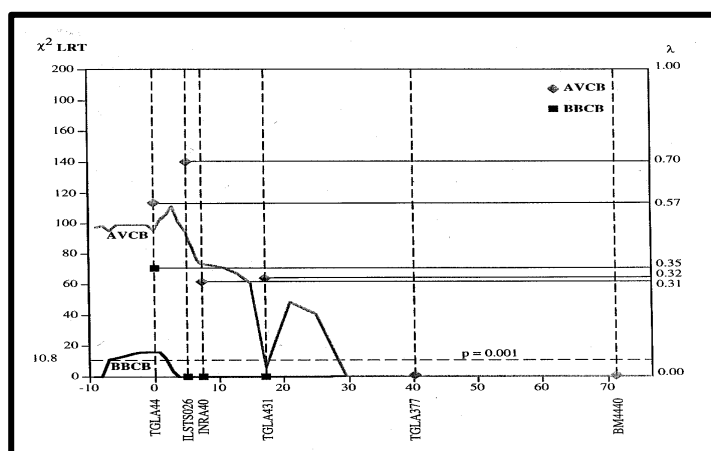
- 1- El fenómeno de la hipertrofia muscular se origina en las razas Blanco-Azul belga y Asturiana por la aparición de un alelo recesivo de un *locus* situado en el cromosoma 2 bovino. Debido a la casi idéntica posición obtenida por los análisis de ligamiento, se planteó la hipótesis de que el mismo gen era el causante de la hipertrofia muscular en estas dos razas.

- 2- El hecho de que en Blanco-Azul belga sólo se observe desequilibrio de ligamiento con el marcador TGLA44, apoya la idea de que el *locus* debía encontrarse en las proximidades de ese marcador. Sin embargo en las familias de raza Asturiana, el desequilibrio de ligamiento se mantiene en una región más larga del cromosoma. Los autores explicaron este fenómeno por la existencia de una posible tasa de recombinación menor en esa zona en la raza Asturiana (por ejemplo no se encontraron recombinantes en el intervalo TGLA44-TGLA431). Además un desequilibrio de ligamiento en un fragmento mayor puede ser la consecuencia de una selección a favor de los animales culones en la raza Asturiana más reciente. En esta zona no se encontraban recombinantes entre los marcadores TGLA44 y TGLA431, pero posiblemente este fragmento más largo en desequilibrio de ligamiento sea consecuencia de un efecto fundador, ya que se utilizaron pocos sementales en la raza Asturiana de los Valles en los años 50.

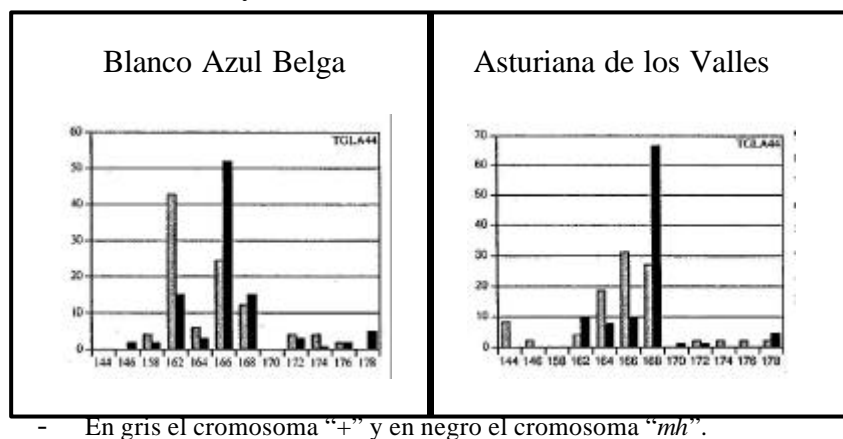
- Figura 7. Localización *locus mh*. Curva de *lodscore* multipunto, análisis de ligamiento (Dunner y col., 1997).



- Figura 8. Localización *locus mh*. Curva multipoint, desequilibrio de ligamiento (Dunner y col., 1997).



- Figura 9. Distribución de la frecuencia alélica del marcador TGLA44 (Dunner y col., 1997).



- 3- El hecho de que el desequilibrio de ligamiento en cada raza sea con un alelo diferente, lo primero que puede indicar es que puedan ser mutaciones diferentes dentro del mismo *locus*. Una misma mutación en razas diferentes implicaría compartir un fragmento de ADN flanqueante al *locus mh*, y por lo tanto un haplotipo idéntico por descendencia. Sin embargo en el marcador más cercano, el alelo asociado al cromosoma *mh* no era el mismo. Podría darse el caso de compartir una misma mutación y que el desequilibrio de ligamiento con TGLA44 hubiera aparecido posteriormente. Esta suposición también estaba reforzada por el hecho de la existencia en la raza Asturiana de los valles de un fragmento mayor de cromosoma en desequilibrio. Si esto es así, el haplotipo idéntico por descendencia flanqueando el *locus mh* debería aparecer al añadir marcadores más cercanos.

En definitiva, se confirma que el *locus* responsable de la hipertrofia muscular bovina es un gen mayor de herencia parcialmente recesiva, localizado en la región centromérica del cromosoma 2 bovino, cercano al marcador TGLA44 con el que se encuentra en desequilibrio de ligamiento. Esta confirmación permitió:

- 1- Iniciar una selección asistida por los marcadores más próximos al *locus mh*, tanto a favor como en contra del carácter culón en estas razas.
- 2- Pensar en una estrategia de búsqueda e identificación del gen responsable.

### 3- Clonado posicional del gen responsable de la hipertrofia muscular bovina.

La estrategia para la identificación de genes responsables de caracteres monogénicos depende de los conocimientos a priori con los que se cuenta. Estos conocimientos se pueden agrupar en dos tipos (Strachan y Read, 1999):

- 1- Físicos. Se refieren a la posibilidad de tener una serie de herramientas relacionadas con el gen, como son:
  - Conocimiento de la localización cromosómica o sub-cromosómica del gen.

- Existencia de anomalías cromosómicas en el cariotipo de los animales afectados. El gen responsable puede verse afectado por la anomalía, bien por que se sitúe dentro de la zona ausente (deleción), duplicada, o translocada, o por que esa translocación se inserte en medio del gen.
- Existencia de genotecas de clones de la región. Las más utilizadas en el caso de genomas de mamíferos son las genotecas de YAC (*Yeast Artificial Chromosome*), BAC (*Bacterial Artificial Chromosome*) y fagos.
- Conocimiento de modelos animales del fenotipo

2- Funcionales. Fundamentalmente se trata del estudio del fenotipo: edad de aparición, tejido/s afectados, estudios macroscópicos e histológicos.

Una vez recogidos todos los conocimientos a priori, existen varias estrategias generales en la búsqueda e identificación de genes implicados en caracteres de interés zootécnico. Todos estos métodos van dirigidos inicialmente a identificar una serie de genes, que pasarán a ser candidatos y que luego deben ser analizados individualmente para demostrar asociaciones con el fenotipo. Se diferencian 4 grandes estrategias:

**a- Clonado Funcional:** Identificación de un gen basándose fundamentalmente en la información sobre la naturaleza bioquímica del carácter, sin ningún conocimiento de la localización cromosómica. De esta manera se ha identificado al gen Procolágeno I aminoproteinasa como responsable de la dermatoparaxia bovina (Nusgens y col. 1992).

**b- Clonado Posicional:** Aislamiento de un gen basándose en su localización subcromosómica, sin usar ninguna información de su patogénesis o su función bioquímica (llamada durante algún tiempo “genética inversa”). Los pasos a seguir son: construir un mapa genético y otro físico lo más detallado posible, identificar genes en esa región estudiada y analizarlos como candidatos, buscando asociaciones con el fenotipo. Suele ser un trabajo demasiado largo y pesado, ya que normalmente hay que reducir el intervalo físico a 1 o 2 Mb. de ADN, antes de iniciar la búsqueda de genes candidatos en la zona.

**c- Clonado de genes candidatos posición-independiente:** Búsqueda de genes candidatos sin tener ningún conocimiento previo de su localización en el genoma. Sólo tiene éxito en el caso de que el fenotipo se parezca mucho a otro en animales o humanos en el que el gen es conocido, o en casos en los que la patogénesis molecular indique que el gen pertenece a una familia de genes conocidos. Utilizando esta estrategia se identificó el gen bovino responsable de la BLAD (*Bovine Leucocyte Adhesion Deficiency*) (Kehrli y col., 1990).

**d- Clonado de candidatos posicionales:** Supone la identificación de genes que, por situarse dentro de una región candidata, se llaman genes candidatos posicionales. Posteriormente estos genes candidatos son analizados para encontrar posibles mutaciones que puedan explicar el fenotipo. Es la estrategia más utilizada debido a que cada vez se están localizando más genes humanos en zonas cromosómicas específicas. Siguiendo una estrategia de clonado de candidatos posicionales en bovinos se han identificado los genes responsables del factor rojo (receptor 1 de melanocortina, Kungland y col., 1995) y de la hipertrofia muscular (miostatina, Grobet y col., 1997).

En general todas estas estrategias conducen a la identificación de uno o varios genes candidatos que posteriormente deben ser analizados individualmente para encontrar evidencias de su asociación con el fenotipo. Un buen gen candidato debe mostrar un patrón de expresión compatible con el carácter (expresión en el mismo tejido que el afectado, mismo patrón temporal, etc.). Una vez encontrado, debe demostrarse que el gen candidato es el *locus* responsable del carácter. Esto se puede hacer de varias maneras:

- Análisis de mutaciones: la identificación de mutaciones en individuos que presentan el carácter y ausentes en el resto, sugiere que se ha identificado el gen correcto.
- Restauración del fenotipo normal *in vivo*. Por ejemplo en situaciones donde la mutación provoca pérdida de función de una proteína, el fenotipo puede ser reversible añadiendo la proteína correcta.
- Producción de un modelo del fenotipo en ratón: Crear un modelo transgénico en ratón por técnicas de *gene targeting* y esperar que el ratón muestre un fenotipo igual o muy parecido.

En un clonado posicional se parte de una región cromosómica candidata que puede ser más o menos larga. Estas situaciones iniciales pueden ser el resultado de varias fuentes de información, lo más común es:

- 1- Pérdida de heterocigosis: identificación de una región subcromosómica que está deleccionada.
- 2- Detección de anormalidades cromosómicas, que pueden afectar a la región candidata del carácter.
- 3- Análisis genético de ligamiento: identificación, en estudios familiares, de un marcador o marcadores en desequilibrio de ligamiento. Es el método más común para definir zonas o regiones candidatas (como en el caso del *locus mh*, Charlier y col., 1995; Dunner y col., 1997).

Debido a los grandes avances del proyecto Genoma Humano, cada vez hay un mayor número de genes cartografiados. Esto conlleva que las estrategias de clonado de candidatos posicionales, que combinan el conocimiento de la posición con la existencia de mapas densos de transcritos, sean hoy en día las más utilizadas para identificar genes (Collins, 1995).

### **3.1 Cartografía física.**

Las técnicas utilizadas para el mapeo físico se dividen en dos clases, dependiendo del nivel de resolución:

- 1- Mapeo físico de baja resolución: cuando se refiere a unidades cuyo tamaño oscila entre una y varias megabases.
- 2- Mapeo físico de alta resolución: cuando el tamaño oscila entre unos cientos de kilobases y un solo nucleótido.

### **3.1.1 Cartografía física de baja resolución:**

Entre las técnicas utilizadas en este tipo de estudios destacan las siguientes:

#### **3.1.1.1 Paneles de células híbridas somáticas:**

Son paneles de células que permiten localizar cromosómicamente cualquier fragmento de ADN. Para construir estos paneles, células de dos especies diferentes son inducidas a fusionarse, bajo ciertas condiciones de cultivo, para formar un *heterokarion*, es decir una célula con dos núcleos diferentes. Cuando estas células entran en mitosis, se eliminan las membranas nucleares y los cromosomas de las dos especies se unen en la placa metafásica. Estos híbridos son inestables y la mayoría de los cromosomas que no son de la especie aceptora (normalmente roedor) no se replican en las sucesivas rondas de división celular y se pierden. Estas pérdidas de material cromosómico son al azar. Como resultado se obtienen una serie de clones, más o menos estables, con contenido completo de ADN de roedor más un fragmento de ADN de la otra especie. La identidad de esos fragmentos cromosómicos que contiene cada clon puede ser analizada por pruebas PCR específicas de un cromosoma, y así construir un panel de células somáticas.

La identificación de los fragmentos cromosómicos en cada uno de los clones permite localizar una secuencia cualquiera de ADN en un cromosoma, ya que dará positivo en todos aquellos clones identificados como positivos para ese cromosoma y negativo en los que no lo son (Abbott y Povey, 1991; Mairal, 1994).

Los paneles de híbridos convencionales normalmente contienen fragmentos de cromosomas distintos, y los resultados pueden ser, a veces, ambiguos. Para resolver este problema también existen paneles de híbridos monocromosómicos. Estos paneles permiten localizar fácilmente una sonda de ADN en un cromosoma (Warburton y col., 1990).

#### **3.1.1.2 Paneles de híbridos de irradiación:**

En esta herramienta, el ADN es fragmentado por irradiación y fusionado con el ADN de célulasceptoras. Así obtenemos una colección de híbridos irradiados.

El análisis por hibridación o PCR de los híbridos de irradiación genera datos que se pueden interpretar estadísticamente para crear un mapa de clones ordenados, ya que cuanto más cerca estén dos secuencias en un cromosoma más baja es la probabilidad de que estén separadas por una rotura debido a la irradiación (Cox y col., 1990). La frecuencia de rotura entre dos marcadores es una medida análoga a la fracción de recombinación meiótica y se representa también como  $\theta$ , pero en este caso varía de 0 a 1, es decir los dos marcadores nunca están separados, o están siempre en fragmentos

diferentes. Debido a que  $\theta$  subestima la distancia entre dos marcadores que están muy apartados en un cromosoma, se utiliza una estima mejor de la distancia, con una función de mapeo análoga a la de Haldane  $D = -\ln(1-\theta)$ , que asume ausencia de interferencia.  $D$  se mide en centiRays (cR) haciendo referencia a la dosis de irradiación utilizada para construir los híbridos. Por ejemplo una distancia  $D = 1\text{cR}_{8000}$  entre dos marcadores indica un 1% de frecuencia de rotura entre ellos después de una exposición de 8.000 radianes de rayos-X (Strachan y Read, 1999).

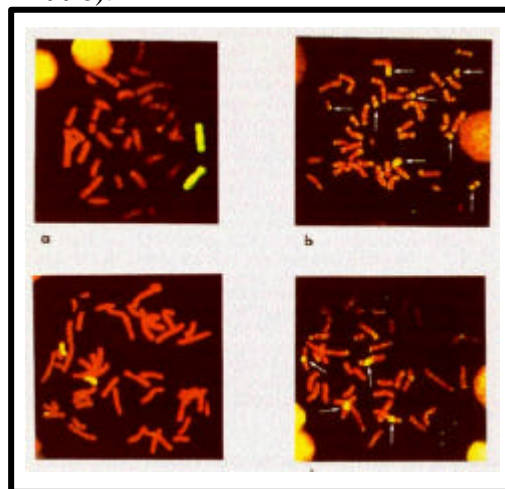
La ventaja principal de estos paneles es que permiten ordenar secuencias únicas de una región definida sin necesidad de buscar un polimorfismo. Estos paneles son la unión ideal entre la cartografía genética y la construcción de un *contig* de clones (YAC o BAC) ordenados.

### 3.1.1.3 FISH (*Fluorescence in situ hybridization*) (*Chromosome painting*).

Esta técnica de hibridación *in situ* permite localizar secuencias únicas, subregiones cromosómicas, cromosomas completos, etc., en células en metafase o interfase. Se puede utilizar para identificar cromosomas, detectar anomalías cromosómicas, o determinar la localización subcromosómica de secuencias específicas. Juega un papel importante en una amplia variedad de áreas de investigación, como la citogenética, diagnóstico prenatal, estudio de tumores, y mapeo genético (Trask, 1991).

En esta técnica, las secuencias de ADN o ARN son marcadas con moléculas *reporter* (biotina, digoxigenina, dinitrophenyl, etc.). La sonda y los cromosomas o el núcleo en metafase/interfase se desnaturalizan, permitiéndose la reasociación de las secuencias complementarias de la sonda y el ADN diana. Después se lava la preparación y se incuba con agentes fluorescentes con afinidad por la molécula *reporter*, de esta forma se debe ver una señal fluorescente en el lugar de la hibridación de la sonda.

- Figura 10. FISH *painting* con sondas humanas y porcinas (Milan y col., 1996).



- *Painting* homólogo y heterólogo con sondas de cromosomas humanos y porcinos.
- Arriba: Sonda del cromosoma 6 porcino sobre metafases de porcino (izquierda) y humano (derecha).
- Abajo: Sonda del cromosoma 19 humano, sobre metafases, humana (izquierda) y porcina (derecha).

El uso simultáneo de varios marcadores coloreados en un protocolo de hibridación hace que su poder de resolución aumente, permitiendo, por ejemplo, el rápido posicionamiento de nuevos clones referido a *loci* conocidos, previamente

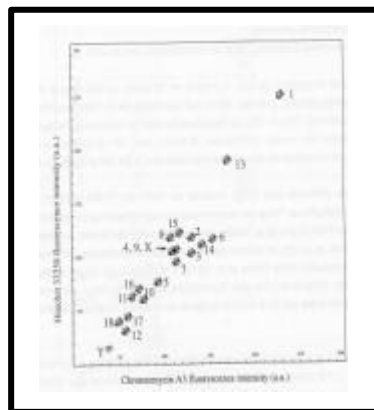
mapeados (van Ommen y col., 1995). Además, se han desarrollado métodos no solo para identificar localizaciones cromosómicas, sino también para el pintado de un cromosoma completo o segmentos de él. Por citometría de flujo o microdissección, los cromosomas aberrantes pueden ser aislados, marcados y utilizados como sondas en protocolos de FISH en cromosomas metafásicos (lo que se ha llamado *reverse painting*). Esto permite identificar reordenamientos y marcadores de cromosomas debido a su patrón de bandeo (Carter y col., 1992; Rack y col., 1993).

#### 3.1.1.4 Citometría de flujo.

La citometría de flujo permite la individualización de células, núcleos, mitocondrias y otras organelas. Más tarde la utilización de colorantes de ADN posibilitó la utilización de la citometría de flujo para analizar los cromosomas (Gray y Langlois 1986, Cram 1990). Esto permite la clasificación de los cromosomas según su contenido en ADN y la obtención en cada especie de un histograma característico llamado Cariotipo de flujo (Riquet, 1996). Pero no es posible realizar cariotipos de flujo en todas las especies. El hombre, mono, el cerdo y el hámster poseen cromosomas que tienen una cantidad específica de ADN cada uno, luego presentan cariotipos de flujo donde cada punto corresponde, la mayoría de las veces, a un solo cromosoma (Gray y Laglois 1986, Grunwald y col. 1986).

En especies como ratón y oveja sólo algunos cromosomas pueden ser individualizados por citometría de flujo en una línea celular normal. En bovino los cromosomas son muy uniformes y las diferencias de tamaño son pequeñas, por lo que la utilización de estas técnicas se ve muy dificultada con respecto a la especie humana. En cambio pueden separarse cromosomas a partir de líneas celulares portadoras de translocaciones robertsonianas, o líneas de híbridos interespecíficos, que a veces permite individualizar algunos picos que corresponden a un solo cromosoma (Baron y col. 1984, Baron y col. 1990, Bahary y col. 1992).

- Figura 11. Identificación de cromosomas de cerdo por Citometría de Flujo (Riquet, 1996).



- Cariotipo de flujo biparamétrico de cromosomas de cerdo. Cromosomas de una línea linfoblastoide de cerdo macho, coloreados con Hoechst 33258 y cromomicina A3. Las dos fluorescencias se recogen de forma separada, lo que permite clasificar los cromosomas dependiendo de su contenido en A-T (eje de ordenadas) y G-C (eje de abcisas).

#### 3.1.2 Cartografía física de alta resolución.

La construcción del mapa físico definitivo, la secuencia nucleotídica completa, no se puede conseguir en el caso de tener una sola molécula de ADN muy larga. Para



obtener un armazón y poder llegar a la secuencia nucleotídica completa, es necesario alinear una serie de clones de fragmentos de ADN, que en conjunto produzcan una buena representación de la secuencia de interés. Para asegurar que la secuencia está completamente representada, esta serie de clones debe estar constituida por segmentos parcialmente solapantes, que forman lo que se llama un *contig* de clones. En principio, el ensamblado del *contig* está facilitado por el hecho de que las genotecas de ADN genómico, están construidas a partir de digestiones parciales del ADN, como parte de una estrategia para maximizar la representación. Debido a estas digestiones parciales los clones individuales normalmente contienen secuencias que solapan parcialmente con algunos otros clones individuales. La construcción de una genoteca se realiza a partir de fragmentos de ADN cuya información de la posición se pierde. Para poder recobrar ese tipo de información pueden utilizarse varios métodos:

#### **3.1.2.1 Paseo cromosómico (*chromosome walking*).**

Esta técnica consiste en establecer un *contig* de clones a partir de puntos de partida fijos. Generalmente se identifican clones de una genoteca utilizando una sonda específica. Los clones positivos obtenidos por hibridación con esta sonda deben contener una secuencia relacionada, incluyendo clones que contienen secuencias que solapan parcialmente. También se puede utilizar una sonda que se localice en el extremo y que esté presente en una sola copia. De esta manera en los nuevos clones positivos se puede hacer la misma operación en rondas sucesivas de hibridaciones. Esto permite la formación de un *contig* a partir de un solo clon por paseo cromosómico bidireccional.

En el caso de utilizar cósmidos se puede simplificar marcando el clon completo para ser utilizado en la hibridación. Sin embargo existe el problema de las secuencias altamente repetitivas que se encuentran en el genoma de mamíferos, y que deben ser suprimidas para evitar señales en la hibridación. Estas técnicas de marcado del clon completo no pueden ser utilizadas en herramientas de almacenado de insertos mayores, como los YAC. En estos casos el mejor remedio es utilizar protocolos dirigidos a la utilización de fragmentos en las extremidades de los YAC (Arnold y Hodgson, 1991; Verhasselt y col., 1992; Libert y col., 1993; Geurts y col., 1994; Weber y col., 1998; Pirottin y col., 1999).

#### **3.1.2.2 Cartografía de STS.**

Un STS, o etiqueta, es un clon corto de ADN (varios cientos de nucleótidos) de secuencia única y conocida sobre la que se pueden diseñar oligonucleótidos y llevar a cabo un protocolo de PCR específico. Si tenemos numerosas etiquetas localizadas en un cromosoma o región cromosómica, se pueden identificar YAC solapantes de la región (por ejemplo porque han sido localizados por FISH) basándose en que contengan o no las etiquetas posicionadas en la región.

Las etiquetas (STS) son herramientas importantes en el mapeo, ya que se puede evidenciar su presencia por PCR. La mayoría de las etiquetas no son polimórficas y deben ser asignadas a localizaciones cromosómicas específicas genotipando paneles de híbridos somáticos. En cambio los microsatélites, etiquetas polimórficas que son usados como marcadores genéticos, pueden ser asignados a localizaciones cromosómicas como

resultado de análisis de ligamiento. Las etiquetas representan una unión entre el mapa genético y físico y proporcionan un lenguaje común para proyectos de mapeo físico, pudiendo ser utilizados para comparar dos *contigs* de clones construidos en laboratorios diferentes (Olson y col., 1989).

### 3.1.2.3 Uso de ADN repetitivo.

El ADN de mamíferos está constituido en una gran parte por ADN repetitivo disperso, como las repeticiones tipo SINE y LINE. Debido a que el espacio entre dos elementos repetitivos puede variar en diferentes regiones cromosómicas, clones que contienen regiones solapantes pueden ser identificados por tener patrones similares como resultado de hibridaciones con sondas de ADN repetitivo (Bellanne-Chantelot y col., 1992).

Las secuencias repetitivas se pueden utilizar también mediante la realización de una PCR con un solo oligonucleótido, que corresponda a un elemento repetitivo muy común. Si un elemento repetitivo es efectivamente muy común, es bastante probable encontrar dos de ellos próximos y en orientación opuesta, así el oligo hibrida y permite la amplificación de la secuencia que se encuentra entre ellos. Esta amplificación, que puede ser una única secuencia o un grupo de ella, se puede marcar y usar como sonda en un protocolo de hibridación para identificar nuevos YAC solapantes (Pirottin y col., 1999).

### 3.1.3 Cartografía física comparada de la región candidata.

La noción de una conservación de ciertos grupos de ligamiento o de sintenia ha ido progresivamente tomando fuerza a partir de los estudios realizados sobre el genoma de diferentes especies. Al principio, las homologías entre los cromosomas se identificaron simplemente por comparación citogenética de los diferentes cariotipos. Los estudios filogénicos realizados por comparación de bandas cromosómicas humanas, de primates y de otros grupos, han aportado elementos nuevos para la comprensión de la especiación (Dutrillaux y col., 1979). Estos trabajos han identificado regiones cromosómicas conservadas entre el hombre y el conejo (Dutrillaux y col., 1980) y el hombre y el gato (Nash y O'Brien, 1982), sin embargo estos estudios únicamente suelen tener éxito cuando se trata de especies poco alejadas. Por este motivo la mayoría de los trabajos para identificar grupos sinténicos conservados entre especies se han basado en hibridaciones *in situ* con sondas genómicas sobre cromosomas en metafase.

Cada vez hay más datos sobre las localizaciones de los genes en las diferentes especies, permitiendo identificar grupos sinténicos conservados entre genomas diferentes (O'Brien y col., 1993). La utilización de estos marcadores de anclaje ha permitido realizar análisis comparativos sistemáticos de los mapas citogenéticos de los diferentes genomas. Dejando a un lado los estudios filogénicos y de evolución, la cartografía comparada es hoy en día una importante estrategia de cartografía de un genoma.

La cartografía de las especies domésticas se apoya en datos de cartografía comparada con los mapas humano y de ratón, para elaborar los mapas regionales de

zonas insuficientemente balizadas en el genoma estudiado. A parte de una simple fuente de marcadores genéticos de una región, la cartografía comparada puede ser utilizada cada vez más, dentro de estrategias de clonado de genes candidatos. Debido a que la zona estudiada contiene un gen implicado en la expresión de un carácter, los genes cartografiados dentro de esa región pueden ser considerados como genes candidatos. Esta estrategia ha sido aplicada con éxito en el estudio del gen *Halotano* en el cerdo (Guérin y col., 1993), y también en la búsqueda de los genes *Booroola* de prolificidad (Lanneluc y col., 1994).

### 3.1.3.1 Cartografía física comparada de la región bovina 2q11-2q21.

La región cromosómica 2q11-2q21 es la región donde se localiza el *locus mh* (Charlier y col., 1995; Fisher y col., 1996; Dunner y col., 1997). En esta misma región se había localizado anteriormente el gen *Col3A1* por hibridación *in situ* (Solinas-Toldo y col., 1995). Además un marcador RFLP de este gen, está ligado al marcador microsatélite TGLA44 ( $\theta = 2\%$ ), el marcador más cercano al *locus mh*, e incluso *Col3A1* ha sido propuesto como candidato posicional del gen de la hipertrofia muscular (Fisher y col., 1996). Estos datos señalaban a la región que flanqueaba a *Col3A1* en humana, HSA2q31-33, como el segmento ortólogo de la región bovina BTA2q11-21.

Para afinar la correspondencia entre los dos segmentos ortólogos, se desarrollaron CATS (*Comparative Anchor Tagged Sequences*, Lyons y col., 1997) para una serie de genes en la vecindad de *Col3A1* en el mapa humano. A parte de *Col3A1* (O'Brien y col., 1993; Lyons y col., 1997) se desarrollaron CATS para los genes *Col5A2*, *INPP1*, *TFP1*, *TTN*, *CHN*, *GAD1*, *CTLA4* y *CD28* (Grobet y col., 1997).

Por otro lado, existía una línea de ratones, que mostraban el fenotipo denominado “compact” (*cmpt*) de forma natural. Estos animales comparados con ratones normales tenían valores de porcentaje de canal superiores (9,4% y 6,8% mayor, en machos y hembras respectivamente) y la relación músculo/hueso era también 1,61% mayor. Además el fenotipo *cmpt* como la hipertrofia muscular bovina también se caracterizaba histológicamente por una hiperplasia (Varga y col., 1997).

El fenotipo *cmpt* mostraba una expresión variable, con una herencia dominante intermedia en machos y casi completamente recesiva en hembras. Por estudios de ligamiento utilizando *pools* selectivos de ADN se encontró un *locus* fuertemente asociado al fenotipo *cmpt* en el cromosoma 1 del ratón. La cartografía fina de la zona, utilizando genotipado individual y análisis de haplotipos, localizó el *locus cmpt* en un intervalo entre los marcadores D1Mit375 y D1Mit21 (Varga y col., 1997). La región del cromosoma 1 del ratón que contenía el *locus Cmpt* era ortóloga a la región humana 2q32-qter (Naylor 1996), es decir los dos *loci mh* y *cmpt* estaban situados en la misma región candidata, con lo cual el gen *cmpt* podía ser el ortólogo del gen *mh* bovino (Varga y col., 1997; Szabo y col., 1998). Debido a estas similitudes entre los dos *loci*, la información disponible del genoma del ratón se añadió en la construcción de la cartografía física de la región candidata bovina (Figura 12).

### 3.1.3.2 Construcción de un *contig* de YAC.

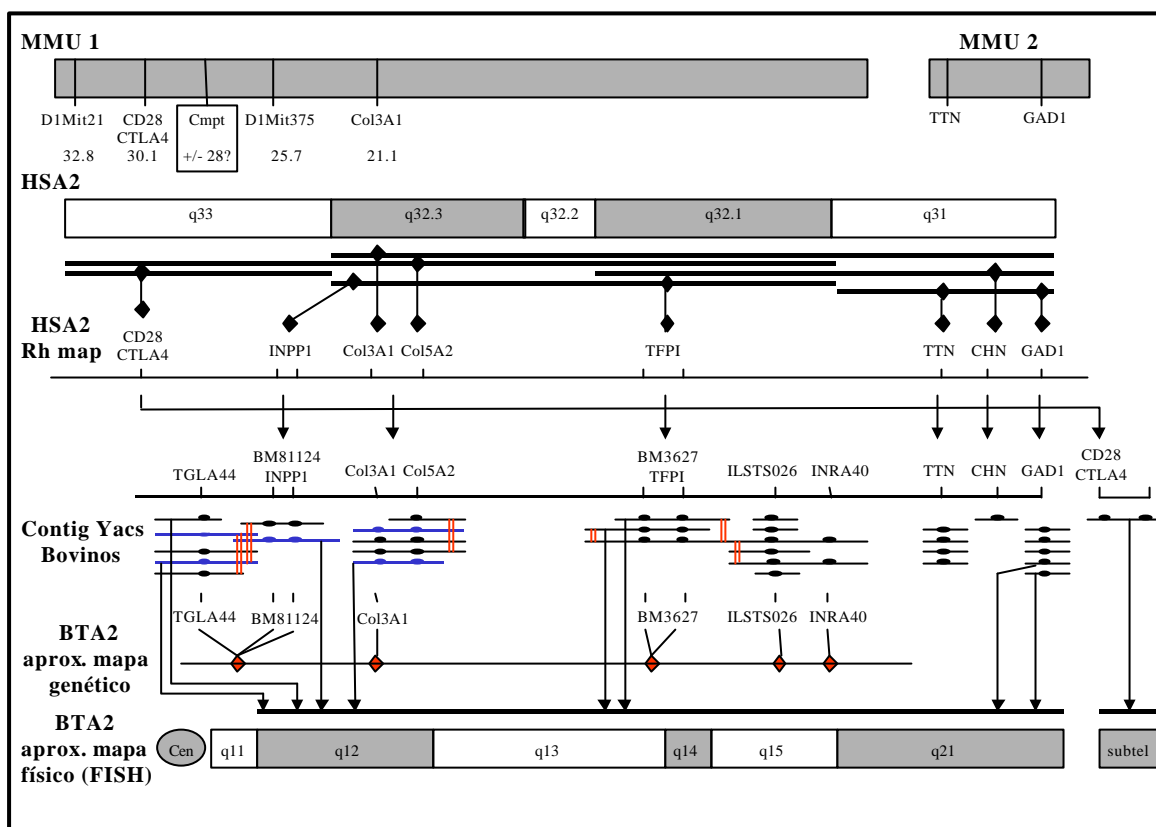
Con los CATS desarrollados, y con los marcadores de tipo microsatélite localizados en la región (TGLA44, BM81124, BM3627, ILSTS026 e INRA40) se analizó una genoteca de YAC de 6 equivalentes de genoma bovino (A. Schoberlein, datos no publicados). Utilizando las etiquetas descritas mediante PCR sobre la genoteca de YAC y también por técnicas de hibridación entre productos de PCR inter-SINE (Pirottin y col., 1999) se encontraron una serie de clones solapantes que se utilizaron para iniciar la construcción de un *contig* de YAC. En la Tabla V se muestran las etiquetas (STS) utilizadas en la construcción del *contig* de YAC de la región candidata.

En la cartografía de la región candidata, se utilizaron los datos procedentes de la cartografía comparada con el genoma del ratón (MM1-MM2, *locus cmpt*), cartografía física de la región ortóloga humana (HSA2q31-2q33 (HSA2)), cartografiado por híbridos de irradiación de los genes localizados en la región (HSARhmap); la aproximación a un mapa bovino de la región BTA2q11-2q21, tanto físico como genético (con los marcadores localizados en la región). Todo estos datos fueron representados sobre el *contig* de YAC de la región candidata que se estaba construyendo y se recogieron de forma gráfica en la Figura 12, que englobaba el conocimiento que se tenía en ese momento de la región candidata del *locus mh*.

- Tabla V. STS en la región BTA2q11-21.

Etiquetas	Nombre del <i>locus</i>	Cartografiado	Referencias
Col3A1	collagen, type III, alpha 1 chain	2q12-q13	Fisher y col.1996; Kappes y col. 1997; Sonstegard y col. 1997.
Col5A2	collagen, type V, alpha-2 chain	2q14-q15	Masabanda y col., 1998.
INPP1	inositol polyphosphate-1-phosphatase	Chr. 2	Grobet y col, 1997.
TFP1	tissue pathway factor inhibitor	2q14-q15	Masabanda y col. 1998.
TTN	titin	2q21	Masabanda y col. 1998.
CHN			
GAD1	glutamic acid decarboxylase	2q21	Masabanda y col. 1998.
CTLA4	Proteína4 asociada a linfocitos-T		
CD28	Antígeno CD28	2q31	Masabanda y col. 1998.
TGLA44		2q12	Kappes y col. 1997.
BM81124		Chr. 2	Stone y col. 1995;Kappes y col. 1997.
BM3627		Chr. 2	Kemp y col 1995; Ma y col. 1996; Kappes y col. 1997.
ILSTS026		Chr. 2	Kemp y col 1995; Ma y col. 1996; Kappes y col. 1997;
INRA40		Chr. 2	Vaiman y col. 1994; Kappes y col. 1997; Slate y col. 1998.

- Figura 12. Cartografía de la región BTA2q11-2q21.



- En azul se representan los YAC elegidos para aislar microsatélites.
- ● YAC positivo a un STS.
- ▮ YAC positivo por hibridación productos de PCR inter-SINE (Pirottin y col., 1999).

### 3.2 Cartografía genética. (ver Capítulo 1).

El *locus mh* se localizaba en la región centromérica del cromosoma 2 bovino, en un intervalo flanqueado por los marcadores tipo microsatélite Tgla44 e Ilsts026. Con la adición de nuevos marcadores a la región candidata se intentaba reducir al máximo posible la región donde se encontraba el *locus*, para poder configurar un mapa físico lo suficientemente pequeño para poderse plantear empezar a buscar genes candidatos.

Se identificaron 5 microsatélites (Bulge18, Bulge20, Bulge23, Bulge27, Bulge28), que permitieron localizar el *locus mh* en un intervalo cromosómico de unos 3cM., entre los microsatélites Bulge18 y Bulge20 (Pirottin y col., 1999).

### 3.3 Identificación de genes candidatos.

El cartografiado fino de la región permitió localizar el *locus mh* en la región BTA2q12, dentro de un intervalo cromosómico de 3 cM. entre los microsatélites Bulge18 y Bulge20, y entre los genes INPP1, y Col3A1. El siguiente paso era la identificación de genes en ese intervalo, que pasarían a ser considerados genes candidatos posicionales.

### 3.3.1 El gen GDF-8 (McPherron y col. 1997).

La familia de factores de crecimiento TGF- $\beta$  está constituida por un gran grupo de factores diferentes que tienen un papel importante, por un lado en la regulación del desarrollo embrionario, y por otro en el mantenimiento de la homeostasis de los tejidos de los animales adultos. En mayo de 1997 aparece publicado en *Nature* el descubrimiento de un nuevo miembro de la familia TGF- $\beta$  (*Transforming Growth Factor- $\beta$* ), GDF-8 (*growth/differentiation factor-8*), que se expresa de forma específica en músculo, tanto en desarrollo como adulto. Durante los primeros estadios de la embriogénesis se expresa en el miotomo de los somitos en desarrollo, mientras que en los últimos estadios y en el animal adulto se expresa en muchos músculos del cuerpo. Para descubrir la función biológica de este nuevo factor de crecimiento, se construyó un alelo nulo del gen mediante estrategias de *gene targeting* en ratones (*Knock-out*). Los animales homocigotos para este alelo nulo mostraban fenotípicamente un aumento de su masa muscular esquelética. Este aumento de masa muscular era debido a una combinación de hiperplasia e hipertrofia de las células musculares. Esto sugirió que el gen GDF-8 actuaba como un regulador negativo del crecimiento muscular esquelético (McPherron y col. 1997).

La secuencia aminoacídica del GDF-8 murino constaba de 376 aminoácidos, y contenía todas las características propias de la familia TGF- $\beta$ : una secuencia señal para la secreción, un lugar diana para el procesamiento proteolítico, una región carboxiterminal con un patrón conservado de residuos de cisteína, nueve exactamente, como los TGF- $\beta$  y las inhibina- $\beta$ . Como los otros miembros de la familia su secuencia estaba conservada entre las diferentes especies, puesto de manifiesto por la técnica de *Southern* (McPherron y col., 1997), que detectó secuencias homólogas en todas las especies de mamíferos examinadas (ratón, rata, humano, mono, conejo, vaca, cerdo, perro), y también en pollo.

También se comprobó la funcionalidad del lugar de procesamiento proteolítico de la secuencia aminoacídica de GDF-8. De igual modo se comprobó su capacidad de formar dímeros, como en el resto de miembros de la familia TGF- $\beta$ , por medio de análisis de tipo *western blot* en condiciones reductoras y no reductoras.

Para determinar los patrones de expresión de GDF-8, se realizaron hibridaciones *in situ* sobre embriones de ratón en diferentes estados de desarrollo. En todos los estados examinados el ARN mensajero de GDF-8 se expresaba en músculo en desarrollo, encontrándose en los primeros estadios en los somitos. GDF-8 continuaba expresándose en el tejido muscular esquelético adulto, y también aparecía en tejido adiposo.

No se conoce bien el papel que GDF-8 juega en procesos como el desarrollo de los rabdomiosarcomas, la hipertrofia inducida por el ejercicio, o la regeneración del tejido después de una necrosis de tejido muscular. Se puede pensar que GDF-8 o sus moléculas inhibitoras pueden ser importantes en el tratamiento de estados degenerativos del músculo, como distrofias musculares, enfermedades neuromusculares o caquexia debida a cáncer.

McPherron y col.(1997), sugirieron la importancia que este gen podía tener en producción animal, teniendo en cuenta que estaba presente en los animales de granja, como cerdos, vacas y pollos.

Por último, basándose en el fenotipo del ratón homocigoto para el alelo nulo, y en la especificidad de expresión en tejido muscular de este gen, se propuso llamar a GDF-8: miostatina (*myostatin*).

### **3.3.2 La miostatina como gen candidato responsable del fenotipo culón en bovino.**

Los trabajos para el clonado posicional del gen se iniciaron con la construcción de un *contig* de YAC, utilizando la cartografía comparada con el genoma humano y del ratón. Además se seleccionaron unos YAC, de los cuales se aisló un grupo de microsatélites que, utilizados en familias donde el gen segregaba, permitirían acortar el intervalo cromosómico que contenía el gen. Los estudios de ligamiento multipunto localizaron este *locus* en el intervalo entre Bulge18 y Bulge20. De acuerdo con la posición relativa de los microsatélites con respecto a los genes INPP1 y Col3A1, deducida de la integración de los mapas humanos y bovinos, se indicó que era muy probable que el gen *mh* se localizara en una región cromosómica flanqueada por los genes INPP1 y Col3A1 (Grobet y col, 1997).

Cuando se publicó en la revista *Nature* el descubrimiento del gen de la miostatina (GDF8), cuyo fenotipo era compatible con el de la hipertrofia muscular, pasó a ser considerado como un candidato del gen culón.

### **3.4 Análisis de mutaciones en el gen candidato.**

Las características del gen de la miostatina lo convertían en un buen candidato para el síndrome de la hipertrofia muscular bovina. El siguiente paso era conocer su posición en el genoma bovino. Los estudios genéticos localizaban el *locus mh* en el cromosoma 2 en un intervalo entre los genes INPP1 y Col3A1. Era necesario confirmar si la miostatina se localizaba en ese intervalo.

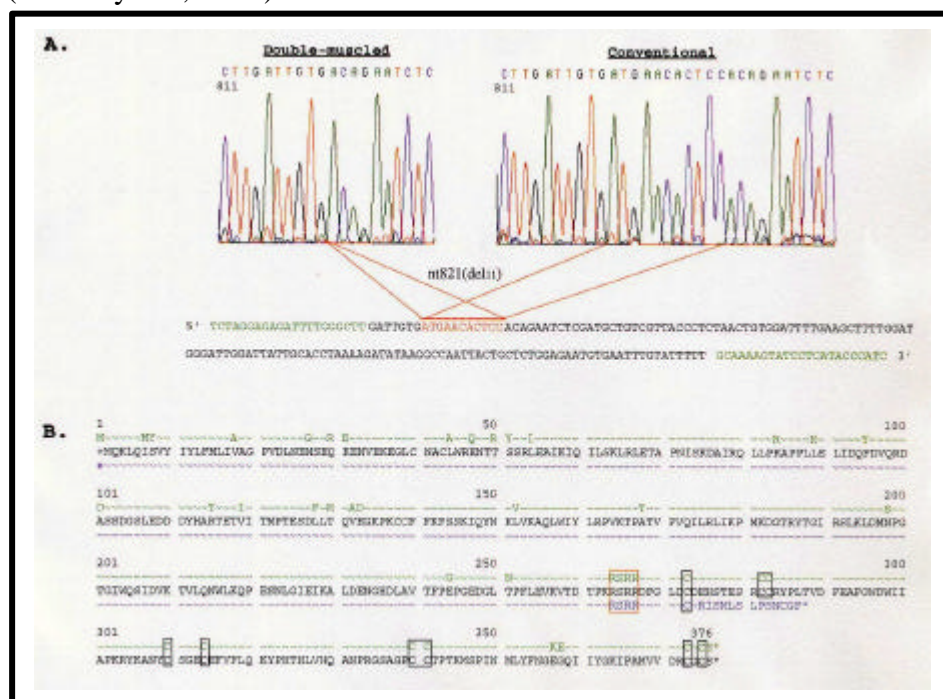
Para localizar la miostatina bovina había que basarse en la conservación de secuencias en genes ortólogos, es decir se quería conocer si el ortólogo humano de la miostatina de ratón se situaba entre INPP1 y Col3A1. Utilizando la secuencia codificante de la miostatina murina se identificaron tres clones de ADN codificante (221299, 300367, 308202) y seis EST (*Expressed Sequence Tags*) (H92027, H92028, N80248, N95327, W07375, W24782). Con estas secuencias se confeccionó una secuencia consenso que representara al ortólogo humano de la miostatina de ratón y se diseñaron oligonucleótidos para amplificar por PCR un fragmento de la miostatina humana. Este protocolo de PCR se utilizó para genotipar un panel de híbridos de radiación (*Genebridge-4*). Esta práctica permitió localizar efectivamente la miostatina humana en el intervalo entre INPP1 y Col3A1, con lo que la miostatina bovina pasó a ser considerado como un importante candidato posicional.

### 3.4.1 Una mutación en el gen de la miostatina es responsable de la aparición del fenotipo culón en la raza Blanco-Azul Belga y Asturiana de los Valles. (Grobet y col., 1997).

Para demostrar si la miostatina estaba asociada al fenotipo culón en bovino, se diseñaron oligonucleótidos basados en las secuencias de ratón y humana para amplificar en su totalidad el ADN codificante bovino a partir de ARN muscular. Se amplificó el ADN codificante a partir de músculo esquelético de dos individuos, supuestamente de fenotipo homocigoto salvaje y homocigoto culón y se secuenciaron las dos hebras del ADN. La secuencia del alelo salvaje mostró una homología del 89 % con la secuencia codificante de la miostatina de ratón, y la proteína una homología de 92 % con la secuencia aminoacídica del ratón.

La secuencia del alelo culón resultó idéntica a la del alelo salvaje excepto por una delección de 11 nucleótidos (Figura 13). Esta delección se sitúa después de la primera cisteína y cambia la fase de lectura de la proteína, lo que provoca la aparición de un codón STOP prematuro. La mutación elimina la parte bioactiva de la proteína, lo que puede ser la causa de la hipertrofia muscular. Siguiendo las normas convencionales de nomenclatura se denominó, a esta mutación, nt821(del11).

- Figura 13. Secuencia de ADN codificante obtenida a partir de un individuo culón y uno normal, mostrando la delección nt821del(11) (Grobet y col., 1997).



- Secuencia de un individuo normal y uno culón, donde se ve la delección de 11pb en el animal culón. En la parte de abajo se muestra la secuencia aminoacídica del ratón (en verde), bovino normal (negro) y culón (azul). Como modelo se utiliza la secuencia bovina normal. El lugar de proteólisis está encuadrado en rojo y las 9 cisteínas conservadas están encuadradas en negro.

Para comprobar que la mutación detectada era la responsable del fenotipo culón se diseñaron oligonucleótidos que flanquearan la delección, para desarrollar un test diagnóstico basado en la PCR. Se genotiparon 100 animales de fenotipo corriente de

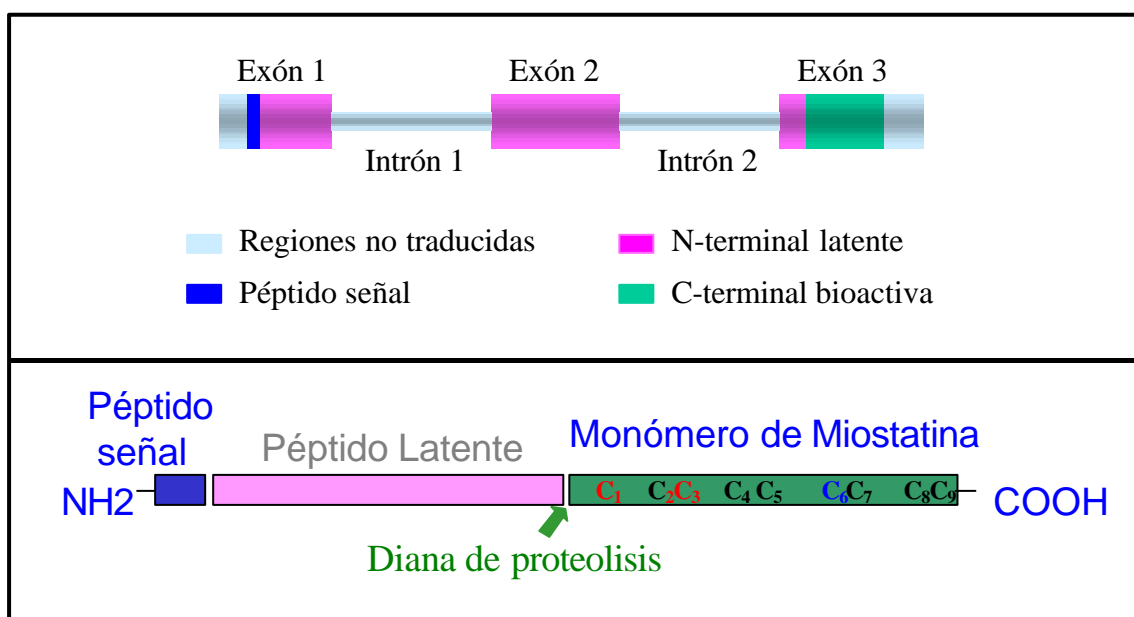


raza Holstein y 96 Jersey. También se genotiparon 104 individuos culones de raza Blanco-Azul belga, 52 de Asturiana de los Valles y 42 Maine-Anjou. El papel del *locus mh* en estas tres razas estaba demostrado (Charlier y col., 1995; Dunner y col., 1997; Grobet y col., 1997). Los 196 individuos de razas lecheras resultaron ser homocigotos para el alelo salvaje. Todos los individuos supuestamente culones de las Razas Blanco-Azul belga y Asturiana de los Valles resultaron ser homocigotos para la mutación nt821(del11). Sin embargo los 42 individuos de la raza Maine-Anjou, fenotípicamente culones, eran portadores del alelo salvaje en homocigosis como los animales de fenotipo corriente de las razas lecheras. Posteriormente se comprobaría que la mutación responsable del fenotipo en esta raza se localizaba en otra zona diferente del gen, y por lo tanto no estaba incluida en este diagnóstico, por lo que estos animales daban un resultado igual a los animales corrientes

### 3.4.1.1 Características básicas de la miostatina bovina.

La miostatina bovina es una proteína de la superfamilia TGF- $\beta$ , que consta de 375 aminoácidos. Como todos los miembros de esta familia se supone que se sintetiza como una molécula precursora larga (pre-pro-miostatina). La secuencia contiene un grupo de aminoácidos hidrofóbicos en la región N-terminal (aminoterminal) que actúan como secuencia señal para la secreción. Después cada molécula debe ser secretada como una forma precursora, que es proteolíticamente procesada en un lugar específico. Tras la proteólisis, queda libre la región C-terminal (carboxiterminal), que es la región en la que los miembros de la familia muestran un cierto grado de homología en la secuencia primaria de aminoácidos, y que contiene nueve residuos de cisteína separados por un espacio característico. Por último, la forma bio-activa de la miostatina está formada por un dímero de dos regiones C-terminal unidas por puentes disulfuro (McPherron y Lee, 1996). El gen de la miostatina está constituido por tres exones separados por dos intrones (McPherron y Lee, 1997; Grobet y col., 1998). En la Figura 14 se muestra de forma gráfica las características de este gen.

- Figura 14. Características generales del gen y la proteína de la miostatina bovina.



### 3.4.2 El fenotipo culón esta causado por una serie alélica de mutaciones disruptivas en el gen de la miostatina bovina. (Grobet y col., 1998).

Una delección de 11 pares de bases (nt821(del11)) en la secuencia del gen de la miostatina, que causa un truncamiento de la parte carboxiterminal de la miostatina era la responsable del fenotipo culón en las razas Blanco-Azul Belga y Asturiana de los Valles (Grobet y col., 1997). Sin embargo en la raza Maine-Anjou, en donde el análisis de ligamiento parecía indicar que era el mismo *locus* el implicado en el fenotipo de hipermuscularidad, no se encontró la delección en los animales analizados (Grobet y col., 1997). Estos datos pueden indicar la existencia de una heterogeneidad alélica en el gen de la miostatina, responsable de la hipertrofia muscular. Esta hipótesis se confirmó con la identificación de una nueva mutación en animales hipermusculados de la raza Piamontesa. En este caso se trata de una sustitución de una cisteína por una tirosina (C313Y) (Kambadur y col., 1997). Para aclarar esta hipótesis de heterogeneidad alélica en el gen de la miostatina o incluso una posible heterogeneidad de *locus*, Grobet y col. (1998) propusieron secuenciar completamente la región codificante del gen de la miostatina en individuos de 10 razas vacunas europeas donde existían individuos hipermusculados y 2 razas de fenotipo convencional utilizadas como secuencia control.

Basándose en la conservación de secuencias en los extremos exón-intrón, se diseñaron oligonucleótidos en los exones 1 y 2 que se suponían flanqueaban los intrones. Con estos oligonucleótidos se llevó a cabo una PCR a partir de un YAC que contenía la miostatina bovina (Y179A3, Grobet y col., 1997). Los productos de la PCR se secuenciaron, y se alinearon con la secuencia conocida del ADN codificante de la miostatina bovina. Tras este alineamiento de secuencias se pudieron diferenciar las uniones exón-intrón predichas, así como los exones y los intrones. Posteriormente se diseñaron protocolos de PCR para amplificar la secuencia codificante completa de la miostatina bovina a partir de ADN genómico y se secuenciaron.

Mediante la secuenciación de los individuos descritos de las diferentes razas europeas utilizadas, se identificaron siete variantes nucleotídicas dentro de la región codificante (Figura 14). Estas mutaciones se agrupan, dependiendo del efecto que tienen sobre la proteína resultante, en tres grupos:

#### 1- Mutaciones que provocan la aparición de un codón Stop prematuro:

- nt821(del11): una delección de 11 pares de bases en la posición 821 después del codón de iniciación, y que provoca la aparición de un codón Stop prematuro. Esta mutación es característica de las razas Blanco-Azul belga y Asturiana de los Valles.
- nt419(del7-ins10): una delección/inserción en la posición 419 a partir del codón de iniciación, donde se reemplazan 7 bases por un grupo de 10 bases. Provoca la aparición de un codón Stop prematuro en la región del péptido N-terminal latente, en la posición aminoacídica 140. Se encuentra en la raza Maine-Anjou.
- Q204X: una transición C→T en el segundo exón en la posición 610, asociada a la aparición de un codón Stop prematuro en el péptido N-

terminal latente en la posición aminoacídica 204. Esta mutación es típica de la raza Charolés.

- E226X: se trata una transversión G→T, en el segundo exón en la posición 676, que conlleva la aparición de un codón Stop prematuro en el péptido N-terminal latente, en la posición aminoacídica 226. Se encuentra en animales de la raza Maine-Anjou.

## 2- Afectan a la estructura secundaria de la proteína.

- C313Y: una transición G→A en la posición 938 que supone la sustitución de una cisteína por una tirosina. Esta cisteína es la quinta de las nueve cisteínas altamente conservadas típicas de la familia TGF-β. Se piensa que está involucrada en la formación de un puente disulfuro intramolecular que estabiliza la conformación tridimensional de la proteína. Esta sustitución es bastante probable que afecte a la estructura y función de la proteína. Esta mutación también ha sido descrita por Kambadur y col. (1997) y McPherron y Lee (1997). Se encuentra en animales de las razas Piamontesa y Gascona.

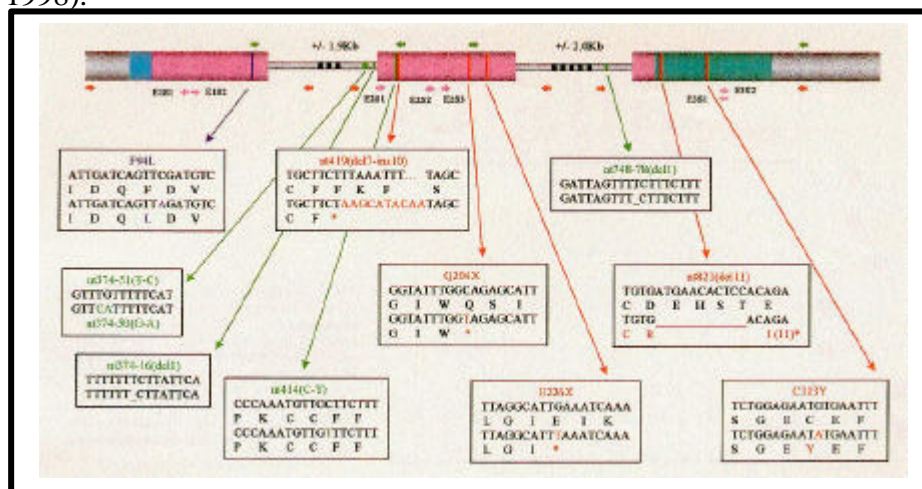
## 3- No afectan a la función de la proteína.

- F94L: una transversión C→A en la posición 282, que conlleva a una sustitución conservativa de una fenilalanina a leucina en el primer exón en la posición aminoacídica 94. Esto ocurre en la parte menos conservada del péptido N-terminal latente. Además en esta misma posición en la miostatina murina aparece una tirosina. Es típica de la raza Limusin.
- Nt414(C-T): es una transición silenciosa C→T, que tiene lugar en la tercera posición del codón 138 (citosina) en el tercer exón. Esta mutación se encuentra en animales de raza Charolés

Además de estos polimorfismos encontrados en la secuencia codificante, se identificaron 4 variantes más en las zonas intrónicas, que probablemente sean polimorfismos neutros. Son los siguientes: en el intrón 1, nt374-51(T-C), nt374-50(G-A), nt374-16(del1); y en el intrón 2, nt748-78(del1).

La mayoría de los animales eran homocigotos para una de las 5 mutaciones disruptivas, o heterocigotos para dos mutaciones disruptivas distintas, lo que era compatible con el modo de herencia recesivo de la hipertrofia muscular en todas estas razas. En dos razas, Limusin y Rubia de Aquitania, no se ha encontrado ninguna mutación disruptiva en la zona codificante de la miostatina lo cual puede indicar dos cosas, o bien que existen mutaciones adicionales del gen de la miostatina fuera de la región codificante, o bien que en el desarrollo exagerado de la musculatura en estos animales no esté involucrado el gen de la miostatina, con lo cual estaríamos delante de fenocopias con un determinismo genético diferente en estas dos razas (Grobet y col., 1998).

- Figura 15. Representación esquemática del gen de la miostatina bovina con la posición de los polimorfismos encontrados (Grobet y col., 1998).



En gris se representa las regiones del ADNc. no traducidas. En azul corresponde al péptido señal para la secreción. En rojo el péptido latente aminoterminal y en verde la parte carboxiterminal bio-activa.

Los polimorfismos encontrados en la secuencia están encuadrados: en rojo se representan los que provocan una miostatina no activa, en azul un polimorfismo, que por situarse en al zona del péptido latente, no parece estar asociado a una modificación de la actividad de la proteína normal y en verde se representan los polimorfismos neutros. Cada polimorfismo se compara con la secuencia control de Holstein.

Aunque el muestreo era limitado, los datos indicaban que algunas mutaciones eran compartidas por varias razas, lo que ponía de manifiesto una cierta migración de los alelos de la miostatina entre diferentes poblaciones, mientras que otras eran específicas de algunas razas. Además, mientras que algunas razas eran muy homogéneas (Asturiana de los Valles o Blanco-Azul belga por ejemplo) otras mostraban una clara heterogeneidad alélica (Maine-Anjou por ejemplo).

Recientemente Miranda y col. han añadido una serie de mutaciones nuevas a esta lista. Utilizando una estrategia de PCR-SSCP se han identificado 7 mutaciones nuevas: 3 mutaciones silenciosas: nt267(A-G), nt324(C-T), nt387(G-A); 2 mutaciones conservativas: D182N, S105C; y 2 polimorfismos intrónicos: nt 747+7(G-A), nt747+11(A-G) (Miranda y col. 2001).

- D182N: cambio aminoacídico que pasa aspártico a asparaguina, no localizado en el péptido bioactivo. Se ha encontrado en las razas Maine d'Anjou y en la línea sintética Inra 95.
- S105C: cambio aminoacídico que no afecta al péptido bioactivo de la miostatina, produce un cambio de serina a cisteína. Se ha encontrado en la raza Parthenaise.

### 3.5 Desarrollo de un test diagnóstico de las mutaciones. Selección asistida por marcadores.

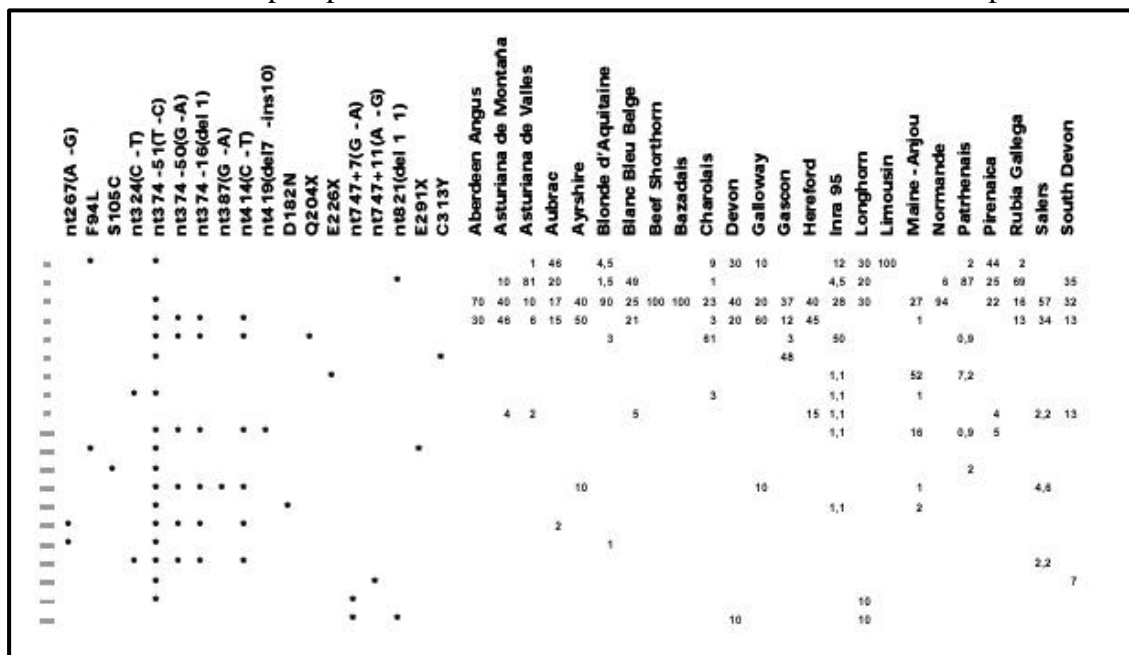
Una vez identificado el gen responsable de la hipertrofia muscular, se abren dos grandes caminos en el campo de la producción animal: por un lado la selección asistida por marcadores, y por otro lado las estrategias de modificación génica o ingeniería genética.

Se han identificado, hasta ahora, 6 mutaciones que afectan a la secuencia codificante de la miostatina bovina.(Grobet y col., 1997; Kambadur y col., 1997; McPherron y Lee 1997, Grobet y col., 1998) causantes del fenotipo de la hipertrofia muscular bovina. En el capítulo 3.4.2 se han revisado 5 de ellas. La última descrita, por el momento, es una transversión G-T en el nucleótido 874, que provoca la aparición prematura de un codón STOP en el dominio carboxiterminal bioactivo. Esta mutación ha sido identificada en animales de la raza Marchigiana (Cappucio y col., 1998). Para seguir estrategias de selección a favor o en contra del fenotipo culón, basta con desarrollar un test diagnóstico de estas mutaciones, para poder identificar a los animales portadores y utilizarlos o no como reproductores dependiendo de la política de selección utilizada.

Karim y col. (2000) y Miranda y col. (2002), han desarrollado test para la detección simultánea de las 6 mutaciones disruptivas descritas. Este protocolo de diagnóstico se basa en la técnica OLA (*Oligonucleotide ligation assay*; Grossman y col., 1998). A parte de este protocolo de diagnóstico conjunto de todas las mutaciones descritas por Grobet y col., 1998, y Cappucio y col., 1998, existen otros protocolos basados en la PCR para la detección de algunos de los alelos de la miostatina (Antoniou y Grosz, 1999; Fahrenkrug y col., 1999).

Además, para evaluar la influencia que ejercen las diferentes mutaciones o polimorfismos del gen de la miostatina sobre el fenotipo se ha desarrollado un estudio en la marco de un proyecto europeo (BIO4-CT98-0421) que consiste en la descripción de los polimorfismos existentes en secuencia. El segundo paso es definir los haplotipos diferentes y estimar sus frecuencias en diferentes razas europeas, además de analizar con detenimiento los fenotipos asociados a cada haplotipo. En la tabla VI se pueden ver los haplotipos de la miostatina encontrados y su frecuencia en diferentes razas europeas (Miranda y col., 2001).

- Tabla VI. Haplotipos de la miostatina bovina en diferentes razas europeas.



- Mutaciones del gen de la miostatina ordenadas en los 20 haplotipos encontrados. A la derecha figuran las frecuencias con las que cada haplotipo ha sido hallado en cada una de las razas comprendidas en el estudio.

## Objetivos.

---

El departamento de Producción Animal de la Facultad de Veterinaria de la U.C.M. de Madrid, junto con las asociaciones de ganaderos ASEAVA y ASEAMO, participa en diferentes proyectos:

- programas de mejora genética de las razas asturianas Asturiana de los Valles y Asturiana de la Montaña.
- comparación de características de crecimiento y desarrollo muscular entre animales homocigotos y heterocigotos para el gen de la hipertrofia muscular frente a los animales libres de dicho gen.
- localización del gen de la hipertrofia muscular y las posibilidades de su utilización tanto en raza pura, como en cruzamiento industrial.

Fruto de estos trabajos y de la colaboración, desde 1995, con investigadores del Departamento de Genética de la Universidad de Lieja (*Department of Genetics, Faculty of Veterinary Medicine, University of Liège*), se publicó, en el año 1997, un artículo que permitió identificar al gen responsable de la hipertrofia muscular presente en la raza Asturiana de los Valles, como el mismo gen que causaba un fenotipo similar en animales de raza Blanco Azul Belga, y que se localizaba en el cromosoma 2 bovino (Dunner y col., 1997).

El hecho de que los análisis de ligamiento realizados en las dos razas localizaran el *locus mh* en el cromosoma 2, permitió iniciar estrategias de selección asistida por marcadores basándose en los alelos asociados con el fenotipo culón. Por otra parte la localización del locus en un segmento cromosómico concreto (2q12-21), abrió la posibilidad del inicio de una estrategia de clonado posicional del gen culón, y plantear el trabajo de esta tesis con la financiación CICYT AGF98-1087.

Tras la puesta de manifiesto del gen responsable de la hipertrofia muscular bovina, la miostatina (McPherron y col., 1997; Grobet y col., 1997), se replantearon los objetivos del trabajo de tesis.

Los objetivos concretos de este trabajo son:

- 1) Contribuir al clonado posicional del gen responsable de la hipertrofia muscular bovina, reduciendo el intervalo cromosómico en el que se encuentra el gen.
- 2) Secuenciar la totalidad del gen de la miostatina bovina y sus regiones flanqueantes.
- 3) Identificar y analizar las posibles regiones reguladoras de la expresión del gen de la miostatina.

## Capítulo 1.

### **Contribución al clonado posicional del gen culón. Reducción del intervalo cromosómico.**

---

Lo primero que se necesita para reducir el tamaño de una región candidata es un conjunto de marcadores polimórficos localizados en la zona. Lo habitual es partir de una serie de clones representativos de la región, como fuente para aislar marcadores. Estos marcadores permitirán construir los mapas genéticos y físicos, y dentro de estos mapas es donde se buscarán los genes candidatos.

La técnica más frecuente para identificar clones de regiones específicas consiste en analizar genotecas de clones, siendo las más utilizadas las de YAC y BAC. Otros métodos para la búsqueda de clones de regiones específicas son: la microdissección de cromosomas o de bandas cromosómicas concretas (Lüdecke y col., 1989), el clonado substractivo (Lisitsyn y col., 1993), o el análisis de genotecas específicas de cromosomas o bandas cromosómicas.

#### **1 Elección de YAC para aislar microsatélites.**

##### **1.1 Descripción de la genoteca bovina de ADN genómico en YAC.** (A. Schoeberlein y col., datos no publicados).

Genoteca de 6 equivalentes de ADN genómico bovino. La genoteca está constituida por 32.256 clones almacenados en medio de cultivo con glicerol a  $-80^{\circ}\text{C}$ , en 336 placas de 96 pocillos (8x12), con un tamaño medio estimado de 548 Kb (Pirottin y col., 1999).

##### **1.2 Organización y análisis de una genoteca de YAC.**

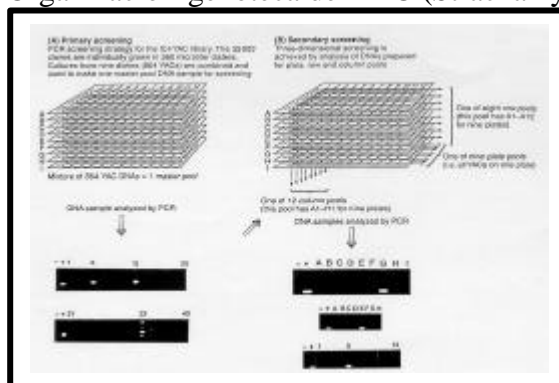
El ADN de cada uno de los clones de la genoteca se organiza en *pooles* para facilitar su análisis por PCR siguiendo un esquema de *pooles* tridimensionales (Libert y col., 1993). Las 336 placas se dividen en 28 grupos de 12 cada uno. El ADN de los clones integrantes de estas 12 placas de cada grupo (96x12), va a constituir un *superpool*. Dentro de cada *superpool*, existen también *pooles* de ADN correspondientes a cada una de las tres dimensiones, placas, filas y columnas. Por un lado se unen los ADN de cada placa (12 *pooles* de 96 cada uno, primera dimensión), de cada fila (8 *pooles* de 144 cada uno, segunda dimensión), y de cada columna (12 *pooles* de 96 cada uno, tercera dimensión).

De esta forma la búsqueda de un clon positivo se realiza con un número mínimo de PCR. La primera PCR se lleva a cabo sobre los *superpooles*. En cada *superpool*



positivo, se hace PCR en los *pooles* de las tres dimensiones (12,8,12) y se debe encontrar, al menos, un positivo por dimensión. Si ese es el caso se cuenta con las tres dimensiones que indican exactamente las coordenadas del clon positivo, es decir el grupo de doce placas, y dentro de éste la placa, la fila y la columna. A veces puede faltar alguna de las dimensiones, y en este caso deben analizarse individualmente todos los clones posibles. De todas formas siempre debe confirmarse el positivo por PCR sobre ADN del clon individual.

- Figura 16. Organización genoteca de YAC (Strachan y Read, 1999).



Para reducir al máximo posible la región cromosómica donde se localizaba el *locus mh* se decidió saturar la región candidata (2q11-13) de marcadores tipo microsatélite. Para ello, se aislaron los marcadores a partir de clones representativos de la región utilizando los YAC situados en el *contig* que se estaba construyendo (ver Figura 12).

Se localizaron algunos STS en la región BTA2q11-2q21 (Figura 12), pero se eligieron como principales puntos de anclaje para la búsqueda de marcadores, por un lado el marcador microsatélite TGLA44 y por otro el gen Col3A1. Por este motivo se decidió empezar a buscar marcadores en los YAC que contenían estos STS, y los YAC de la zona que quedaba entre ellos. Los YAC se eligieron después de confirmarse por FISH que hibridaban en la zona correcta y por tener el menor grado de quimerismo posible. Se utilizaron más de un YAC de cada STS para evitar una falta de representación de la zona.

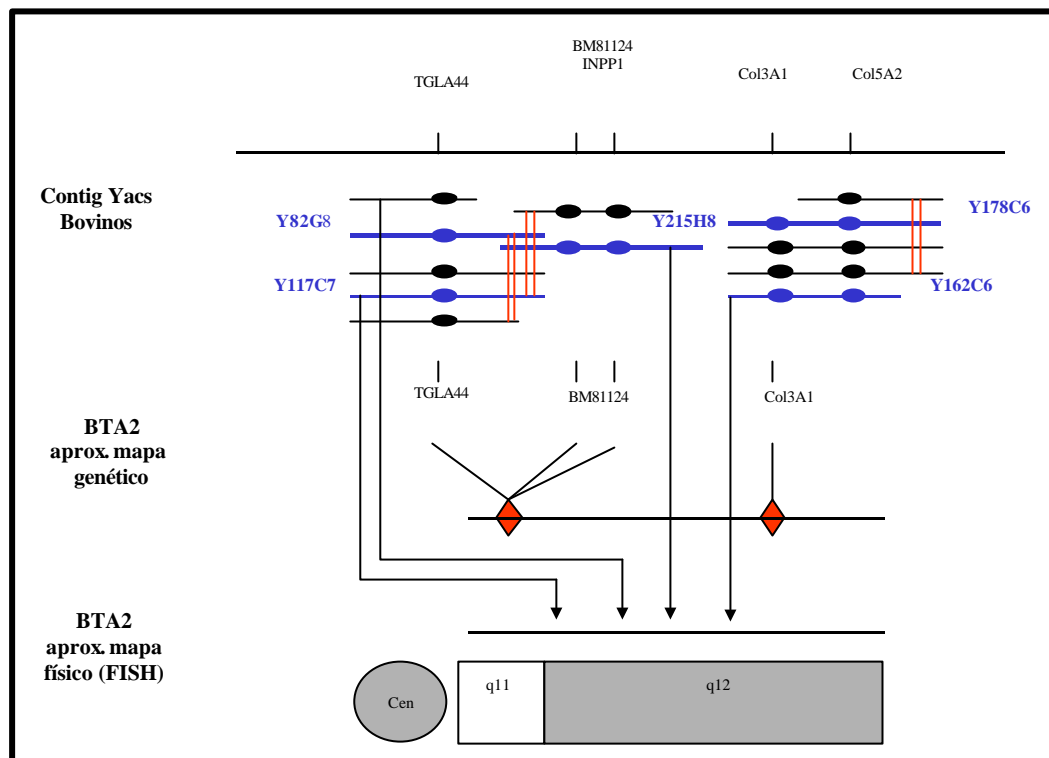
En la Figura 17 se aprecia la localización de los YAC elegidos. Este grupo de YAC comprende una zona del mapa físico flanqueada por las etiquetas TGLA44 y Col5A2, e incluye YAC positivos a INPP1, BM81124 y Col3A1, además de las dos flanqueantes TGLA44 y Col5A2 (Figura 17, Tabla VII).

- Tabla VII. YAC elegidos para aislar microsatélites.

STS positivo	YAC	Tamaño*	Observaciones
TGLA44	Y117C7	350 Kb	co-migración Cr. 3 de levadura
	Y82G8	350 Kb	Quimérico, co-migr.Cr.3 levadura
BM81124	Y215H8	440 Kb	Co-migrac. Cr. 9 levadura
+INPP1			
Col3A1	Y162C6	590 Kb	Quimérico
Col5A2	Y178C6	900 Kb	Quimérico

\* Los tamaños de los YAC, se estiman a partir de la migración en campo pulsado y su comparación con el tamaño conocido de los cromosomas de levadura.

- Figura 17. Detalle de los YAC elegidos (en azul) para aislar microsatélites.



- En azul YAC elegidos para aislar microsatélites.
- ● Positivos a los STS por PCR.
- || YAC solapantes por protocolos de hibridación inter-SINE (Pirottin y col., 1999)

## 2 Protocolo general de aislamiento de microsatélites.

El tipo de microsatélite más abundante en el genoma de los mamíferos es la repetición CA (1 microsatélite cada 25 o 100.000 pb., Stallings y col., 1991; Kaukinen y Varvio, 1992; Kostia y col., 1997), por eso se buscaron microsatélites utilizando la sonda (CA)<sub>10</sub>. Además de éstos, también se buscaron positivos a una sonda (CT)<sub>10</sub>. Para el aislamiento de los microsatélites se siguieron los siguientes pasos:

### 2.1. Aislamiento del YAC.

Lo primero fue separar los YAC elegidos del resto del genoma de la levadura. Para identificar y separar el cromosoma artificial de los cromosomas de levadura se utilizó un protocolo de electroforesis en campo pulsado (Gardiner y col., 1986; Crete y col., 1991).

#### 2.1.1 Electroforesis en Campo Pulsado.

La migración en campo pulsado permite la separación de fragmentos de ADN desde cientos a miles de kilobases. Se utiliza para separar el cromosoma artificial (YAC) del resto de cromosomas de la cepa de cultivo de *S. cerevisiae*.

Día 1: Siembra de los clones elegidos:

- Sembrar 2 tubos con 5ml. de medio AHC, a partir de *stock* glicerol de cada clon.
- Incubar 2 o 3 días a 30° C. con agitación a 200 r.p.m.

Día 2: Concentración y mantenimiento del ADN de levadura en tacos de agarosa. En este segundo día se preparan estos tacos de agarosa.

- Preparación de los tacos de agarosa:
- Lavar los moldes para hacer los tacos de agarosa, sumergiéndolos 30' en HCl 10%.
- Aclararlos con H<sub>2</sub>O UP.
- Centrifugar el medio de cultivo AHC después de dos días a 30° C. 5' a 2000 r.p.m. y eliminar el sobrenadante.
- Resuspender el *pellet* con 2 ml. de EDTA 50mM., pH=8 (si es necesario, utilizar puntas cortadas para evitar romper las hebras de ADN)
- Centrifugar 5' a 2000 r.p.m. y eliminar el sobrenadante.
- Resuspender el *pellet* con 1ml. de SPE. Transferir a un tubo de 1,5 ml.  
Se pueden unir los *pellet* de los dos tubos de un mismo clon.
- Centrifugar 3' a 4000 r.p.m., eliminar sobrenadante.
- Resuspender el *pellet* con 120 µl. de solución de *Spheroblasting*.
- Incubar 1 h. a 37°C.
- Preparar una solución de Agarosa Seaplaque GTG al 1,5 % en EDTA 125mM., pH=8.
- Alicuotar 140 µl. de la agarosa en un tubo de 1,5 ml. y mantener a 50° C., para que se mantenga en estado líquido.
- Incubar las muestras una a una 30'' a 50° C.
- Mezclar la muestra con la solución de agarosa en EDTA caliente.
- Homogeneizar un poco con la pipeta y llenar el molde de los tacos con la agarosa.
- Dejar que se enfríen, y solidifique la mezcla de agarosa con el cultivo de la levadura.
- Poner 2,5 ml. de ETM en placas de 6 pocillos (6 *well-plates*).
- Meter los tacos en las placas con ETM.
- Sellar las placas con parafilm e incubar una noche en ETM a 37° C. con agitación ligera.

Día 3: Lavado de los tacos para su almacenamiento.

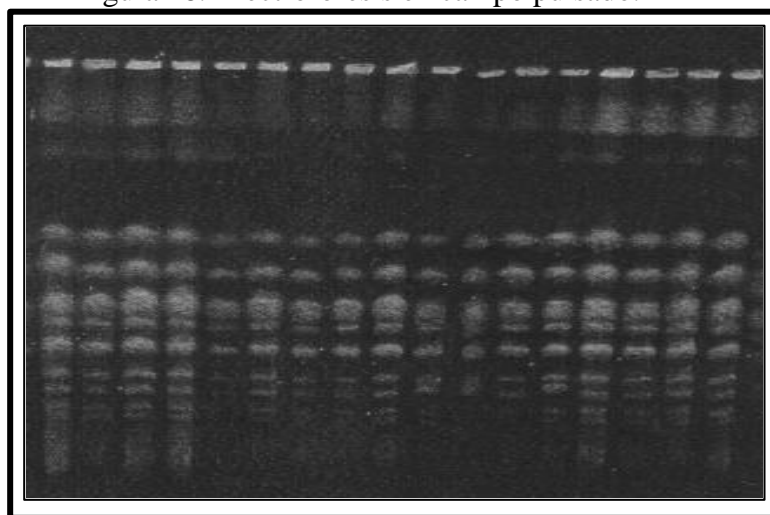
- Aspirar la solución ETM de los pocillos, con cuidado de no dañar los tacos.
- Añadir 4 ml. de LDS.
- Incubar 1 h. a 37° C. con agitación ligera.
- Aspirar la solución LDS.
- Reemplazar por LDS nueva.
- Incubar 1 h. a 37°C. con agitación ligera.
- Aspirar solución LDS.
- Lavar 3 veces con 5ml. de T<sub>10</sub>E<sub>1</sub> pH=8, 30' por lavado, a temperatura ambiente.
- Conservar los tacos en T<sub>10</sub>E<sub>1</sub> pH=8 a 4°C.

Día 4: Electroforesis. En este tipo de electroforesis se van a separar los cromosomas de la levadura que están concentrados y almacenados en los tacos de agarosa preparados los días anteriores.

- Aspirar el  $T_{10}E_1$  de los pocillos.
- Reemplazarlo por 5ml. TBE 0,5x.
- Lavar en esta solución al menos 4 o 5 h.
- Preparar un gel de agarosa *Seaplaque* 1% en TBE 0,5 %.
- Mantenerlo en un baño-maría a 50° C.
- Llenar el campo pulsado con 2 l. de TBE 0,5x
- Hacerlo funcionar al menos 1 h., hasta que la temperatura sea de 13° C.
- Colocar una banda cortada de cada taco de agarosa contra un diente del peine, depositar una gota de agarosa caliente en la base del peine, y dejar enfriar.
- Completar el gel con agarosa.
- Colocar el gel en el campo pulsado para la migración, bajo las siguientes condiciones:  
     6 V/cm, 14° C.  
     *run time*: 24 h.  
     *Initial switch time* 60''  
     *final switch time* 140''

Normalmente es imposible diferenciar el YAC del resto de cromosomas de levadura sin hacer una hibridación con ADN genómico, y puede darse el caso de que el YAC, debido a su tamaño, migre junto con un cromosoma propio de la levadura

- Figura 18. Electroforesis en campo pulsado.



- Gel de agarosa teñido con Bromuro de etidio donde se ve un patrón típico de migración en campo pulsado de los 16 cromosomas de levadura.

### 2.1.2 Transferencia del ADN a una membrana de nailon (*Southern blot*).

Primero se trata el gel para fragmentar el ADN y facilitar que sea transferido a la membrana. Se deja el gel durante unos 5' sobre la lámpara de ultravioletas. Luego se sumerge en HCl 0,2 N. durante 10', se lava con H<sub>2</sub>O destilada varias veces, y se equilibra con NaOH 0,4 M. antes de la transferencia. Se corta una membrana del

tamaño del gel, se sumerge en H<sub>2</sub>O, se saca, se elimina el exceso de H<sub>2</sub>O y se introduce en NaOH 0,4 M.

Se coloca el gel sobre unas hojas de papel 3MM empapados en NaOH 0,4M. y en contacto con una solución de NaOH 0,4M., sobre el gel la membrana y otro papel 3MM del mismo tamaño de la membrana. Encima se colocan papeles absorbentes y peso, se deja la transferencia durante 12 h.

Transcurrido ese tiempo se desmonta todo. Se lava la membrana en 2x SSC durante 5', y se deja secar sobre papel 3MM. Para verificar si todo el ADN ha sido transferido a la membrana, se recupera el gel y se sumerge en una solución 6x SSC para renaturalizar el ADN que pueda quedar, se tiñe con Bromuro de Etidio y se observa en la lámpara de ultravioleta.

### 2.1.3. Hibridación.

Para identificar el cromosoma, se hibridó la membrana con una sonda específica de ADN genómico bovino. La sonda se marcó siguiendo el protocolo de marcado con oligos (*oligolabeling*), (*Random Primers DNA Labeling System*, GIBCOBRL, *Life Technologies*).

Es un protocolo que permite marcar radiativamente un fragmento de ADN en el curso de una extensión a partir de un conjunto de hexámeros utilizados como cebadores. La enzima polimerasa *Klenow* lleva a cabo esta reacción. En la mezcla se incluyen para la extensión deoxinucleósidos trifosfato en 4 soluciones separadas, dATP, dGTP, dTTP, dCTP. Uno de ellos, normalmente dCTP, está marcado radiativamente en su primer grupo fosfato ([ $\alpha$ -<sup>32</sup>P]dCTP). La reacción se inicia a partir de los hexámeros cebadores distribuidos al azar por toda la molécula molde, se van añadiendo los deoxinucleósidos trifosfato y conforme se incluyen [ $\alpha$ -<sup>32</sup>P]dCTP, se va marcando radiativamente la molécula. La reacción de marcado se lleva a cabo a temperatura ambiente durante 1 h., aunque tiempos más largos de incubación pueden resultar en más actividad específica de la sonda.

Protocolo para marcar 25 ng. de ADN:

Desnaturalizar 100°C. 5'

Guardar en hielo, y añadir:

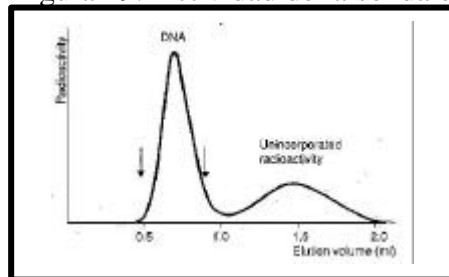
dATP	2 $\mu$ l.
dGTP	2 $\mu$ l.
dTTP	2 $\mu$ l.
<i>Random primer Mix</i>	15 $\mu$ l.
[ $\alpha$ - <sup>32</sup> P]dCTP	5 $\mu$ l.
H <sub>2</sub> O	hasta 49 $\mu$ l.
<i>Klenow</i>	1 $\mu$ l.

Incubar a 25°C. (temperatura ambiente) 1 h.

- Purificación de las sondas en columnas de *Sephadex* (*NICK<sup>TM</sup> Column*, Pharmacia Biotech).

Este sistema de purificación consta de una columna con una matriz de *sephadex* por la que se hace pasar la sonda recién marcada, para separarla de los nucleótidos no incorporados. El uso de este tipo de columnas no está recomendado para purificaciones de oligos menores de 20 bases. El protocolo consta de: Introducción en la columna de la muestra después de haber equilibrado el gel de *sephadex*. Una vez añadida la muestra se añaden 400  $\mu$ l. de tampón (TE). Una vez haya salido el tampón se vuelve a añadir otros 400  $\mu$ l. y se recoge la sonda purificada. A continuación empezarían a salir los nucleótidos radiactivos no incorporados.

- Figura 19. Actividad de la sonda dependiendo del volumen de elución.



- En ordenadas se representa el volumen de elución y en abcisas la actividad de la sonda. Las moléculas marcadas radiactivamente corresponden a la zona delimitada por las dos flechas.

Cuando el cromosoma artificial se ha identificado (es el único que se marca con la sonda de genómico bovino), en una siguiente electroforesis en campo pulsado se recorta la banda que corresponde al YAC y se extrae el ADN

- Colorear con Br.Eth. las bandas laterales.
- Lavar con agua destilada.
- Sobre la lámpara de ultravioleta, identificar los YAC y marcarlos sobre las bandas laterales coloreadas.
- Reconstruir el gel, y recortar las bandas deseadas.
- Equilibrar la banda de agarosa en:
  - EDTA 5mM.
  - NaCl 100mM.
- Mantener a temperatura ambiente con agitación. Repetirlo 2 veces.
- Fundir la agarosa en baño maría a 68°C. durante 10-15'
- Mantener a 37°C. 5-10'
- Añadir agarosa (10u./ $\mu$ l.) para llegar a una concentración final de 50 u./ml.
- Mezclar muy bien, con pipeta si fuera necesario.
- Incubar de 3 a 4 h. a 37°C.
- Extracción orgánica fenol/cloroformo.
- Precipitar al etanol con NaCl 0,3 M.
- Resuspender en 20  $\mu$ l. T<sub>10</sub>E<sub>1</sub>.
- Estimar la concentración en gel de agarosa.

## 2.2 Identificación de microsatélites.

### 2.2.1 Fragmentación del YAC.

Una vez aislado el cromosoma artificial, y extraído su ADN, se digirió totalmente con la enzima *MboI*. La digestión se realizó bajo condiciones estándar a 37°C. con 1 unidad de enzima de restricción por cada µg. de ADN, y en el tampón indicado por el comercial a concentración 1x, durante 2 a 5 h. La reacción se paró inactivando la enzima a 94°C. durante 10'.

Esta digestión produjo fragmentos de tamaño entre 400 y 800 pb., idóneos para su clonado en un plásmido PUC18. Este tamaño permitía además, que en una sola ronda de secuenciación se obtuvieran las secuencias completas de los insertos.

Los fragmentos de la digestión *MboI* se ligaron en el plásmido PUC18/*BamHI*/BAP. Se utilizaron dos relaciones vector/inserto diferentes: 5/1 y 1/1. En 20 µl. de volumen final se mezclan: 50 ng. de plásmido, una cantidad equimolar de ADN digerido, y/o 5 veces menos, dependiendo de la relación vector/inserto utilizado, tampón de ligado 1x, y se incuba con 0,5u. de T4 ADN ligasa durante 2 h. a 16°C.

Controles de ligado:

- a) PUC18/*BamHI*/BAP sin ligasa.
- b) PUC18/*BamHI*/BAP + T4 DNA ligasa.

Los controles se hacen para conocer: a) la cantidad de plásmido no digerido, ya que es capaz transformar a las bacterias, y b) cantidad de plásmido o bien no digerido o digerido pero no defosforilado ya que es capaz de religarse. Con estos dos controles, estimamos la cantidad de clones no recombinantes, así como las condiciones del plásmido utilizado.

Los plásmidos se transformaron en células MAX *Efficiency* DH5-α<sup>TM</sup> *Competent Cells*, y se sembraron en placas de LB-agar.

Protocolo de transformación. En condiciones de esterilidad:

- Descongelar las bacterias sobre hielo.
- Identificar los tubos de transformación (Falcon 2059) y ponerlos en hielo
- Mezclar suavemente las bacterias
- Pipetear 100 µl. para los ligados  
50 µl. para los controles (ligado y transformación)
- Añadir a las bacterias competentes descongeladas, en cada caso:  
1 µl. de ligado  
1 µl. de control de ligado  
2,5 µl. de control de transformación PUC19.
- Mezclar suavemente.
- Incubar 30' en hielo.
- 45 '' en baño María a 42° C.
- 2' en hielo.
- Añadir 0,9 ml. de SOC a temperatura ambiente.
- Incubar 1 h. a 37° C. con agitación (225 r.p.m.).

- Control de transformación:

Añadir a las bacterias competentes un plásmido control, del que se conoce su eficacia de transformación, es decir cuantas colonias transformadas por  $\mu\text{g.}$  de plásmido. Si el resultado de la siembra se encuentra por debajo de la eficacia preconizada del plásmido control, indica algún problema en el protocolo de transformación.

### 2.2.2 Selección de clones positivos.

A partir de las siembras en medio sólido se hicieron réplicas en membrana de Nailon (C/P lift® Membranes, BIORAD). Se sembraron unos 5.000 clones por placa de Petri de 15 cm. de diámetro, ya que esta densidad permite tener un gran número de clones por placa y, si están distribuidos homogéneamente, no hay demasiada confluencia entre las colonias. Esto permite aislar individualmente la colonia positiva en una segunda hibridación.

Para utilizar un método colorimétrico de selección de recombinantes basado en el test de  $\alpha$ -complementación de la  $\beta$ -galactosidasa, se añaden 40  $\mu\text{l.}$  de X-gal + 4  $\mu\text{l.}$  de IPTG por placa.

- Réplicas de placas de clones bacterianos.

Las membranas (BioRad) deben manipularse siempre con pinzas. Previamente hay que enfriar las placas de cultivo a 4° C., durante al menos 1 h.

- Depositar la membrana sobre una placa de LB-agar vacía.
- Una vez humedecida, sacarla y depositarla sobre la placa sembrada.
- Dejarla durante 2 minutos, durante los cuales se marca la membrana con una aguja, con una serie de señales no simétricas.
- Levantar la membrana y ponerla al revés (con las colonias hacia arriba) sobre una placa vacía.
- Incubar 1h. a 37° C. ambas placas.
- Tratamiento de las membranas: pasar las membranas, con la cara de las colonias hacia arriba, por un papel 3MM (Whatman™ 3MM paper) humedecido en las siguientes soluciones, para desnaturalizar, y fijar el ADN:
- Desnaturalización: 5'
- Neutralización: 5'
- Lavar las membranas 1' en solución SSC 2x
- Secar en papel 3MM.
- Fijar el ADN a la membrana por incubación de 2 h. a 80° C.

\*Las membranas pueden usarse inmediatamente o guardarse formando un bocado entre hojas de papel 3MM, a temperatura ambiente.

Estas membranas se hibridaron con las sondas (CA)<sub>10</sub> y (CT)<sub>10</sub> marcadas radiactivamente según el protocolo de marcado en 3'.

Es un protocolo que se basa en el uso de la enzima terminal deoxinucleótido transferasa, que cataliza la unión de dNTP o nucleótidos simples al extremo 3' -OH de



una molécula de ADN monocatenaria o bicatenaria. Se utiliza para marcar radiactivamente moléculas pequeñas.

Protocolo para marcar 20 pmoles de sonda. ( Amersam, Life Science):

20 pmoles ADN  
10 µl. Buffer 5x  
5 µl. [α-<sup>32</sup>P]dCTP  
4 µl. *Terminal transferasa* ( 36 u.)  
11 µl. H<sub>2</sub>O

Purificar la sonda siguiendo el protocolo en columnas de *Sephadex (NICK<sup>TM</sup> Column*, Pharmacia Biotech; ver página 43).

Tras la primera ronda de hibridación, aparecieron 26 clones candidatos positivos a las sondas (CA)<sub>10</sub> y (CT)<sub>10</sub>. De éstos, tras una segunda hibridación, 12 clones resultaron de nuevo positivos y se eligieron como candidatos para secuenciar. Los clones positivos se sembraron en medio líquido, y tras una noche de cultivo a 37° C. se extrajo su ADN plasmídico (protocolo QIAwell 96 *Ultra Plasmid Kit*, Qiagen), y se conservaron en medio de cultivo con glicerol. Para estimar su tamaño se llevó a cabo una PCR con los oligonucleótidos universales *M13-forward* y *M13-reverse* (Tabla VIII).

### 2.3 Secuenciación de los clones positivos.

Los clones positivos se secuenciaron utilizando los oligonucleótidos universales del plásmido PUC18, *M13-forward* y *M13-reverse*

El método utilizado de secuenciación es la secuenciación cíclica usando terminadores marcados con un fluorocromo diferente cada uno (ABI *PRISM<sup>tm</sup> Dye Terminator Cycle Sequencing, Ready reaction kit*).

El protocolo es el siguiente:

<i>Terminator Premix</i>	8 µl.
ADN (150 ng./ µl.)	2 µl.
<i>Primer</i> (3,2 pmol/µl.)	1 µl.
H <sub>2</sub> O UP	9 µl.
Volumen total	20 µl.

El programa de secuenciación se lleva a cabo en un termociclador P.E 9600 (*GeneAmp PCR System 9600*, Perkin Elmer) y consta de 25 ciclos como el siguiente:

96° C. 10''  
50° C. 5''  
60° C. 4'

Precipitación al etanol: Se realiza una purificación de los productos de extensión para eliminar el exceso de terminadores que no han sido utilizados en la reacción, es

decir no incorporados durante la extensión. Esto puede conllevar una pequeña pérdida de los fragmentos de menor tamaño.

En un tubo añadir:

- 2 µl. Acetato sódico 3M., pH=5,2.
- 50 µl. Etanol 95%.
- Transferir los 20 µl. de la reacción de secuenciación.
- Vortex y mantener en hielo 10'.
- Centrifugar a 13.000 r.p.m. 30'.
- Eliminar etanol.
- Lavar con etanol 70%, centrifugar 10' a 13.000 r.p.m.
- Eliminar todo el etanol.
- Secar.
- Utilizar o guardar liofilizado a 4° C.
  
- Preparación del gel de secuenciación. (Secuenciador ABI 373, Perkin Elmer).

Las características utilizadas para un gel de secuenciación en un aparato de tipo ABI 373 de Perkin-Elmer fueron:

Preparar TBE 10x (con productos ultrapuros). Cuando esté bien disuelto, filtrar el contenido. Conservarlo como máximo 15 días a 4°C.

Preparar también una solución 40 % de acrilamida/bisacrilamida (19:1). Se puede conservar hasta 1 mes a 4°C y protegida de la luz.

Diluir 50 g. urea en 30 ml. de H<sub>2</sub>O ultrapura. Para que se diluya bien hay que calentar la solución. Luego añadir 15 ml. de la solución Acrilamida/bisacrilamida 40%, (protegerlo de la luz con un papel aluminio). Mezclarlo con 1-2 g. de resina para eliminar los iones en solución. Dejarlo de 10 a 15 minutos y filtrar.

Añadir 10 ml. de TBE 10x filtrado y ajustar a 100 ml. con H<sub>2</sub>O ultrapura. Se puede guardar a 4°C. durante un mes protegido de la luz.

- Gel de secuenciación:
- 50 ml. de la solución acrilamida/urea 6%, filtrar y desgasificar durante 10'.
- + 250 µl. 10% persulfato amónico.
- + 22 µl. TEMED.
- rellenar el gel y dejar polimerizar 2 o 3 h. a temperatura ambiente.

- Preparación de las muestras:

Resuspender las muestras en 4,5 µl. de tampón de carga FELB (*Formamida EDTA Loading Buffer*). Hacer un *vortex* durante unos segundos y hacer un pico de centrifuga. Repetir varias veces esta operación.

En la Tabla VIII se muestran las características de los clones aislados: el número de clon, el STS al que el YAC de origen es positivo, tipo de repetición encontrada, el tamaño obtenido en la PCR (*M13-forward* y *M13-reverse*), y si su origen es de levadura o no. Para identificar el origen del clon, se realizó un análisis de homología la secuencia con las bases de datos del *GenBank* utilizando el programa *FastA* del grupo de programas GCG. El programa *FastA* se utiliza para buscar similitudes entre una secuencia dada y un grupo de secuencias de la misma clase (ácidos nucleicos o proteínas) que la secuencia dada. El programa busca estas similitudes utilizando el método de Pearson y Lipman (Pearson y Lipman, 1988) De esta forma se identificaron algunos clones (ver Tabla VIII) como secuencias de levadura (por su alta homología con secuencias propias del genoma de *Sacharomices cerevisiae*)

Posteriormente se analizaron las secuencias obtenidas, primero para obtener una secuencia consenso de cada clon (con la ayuda del grupo de programas *Phred/Phrap/Consed package*; Ewing y col., 1998; Ewing y Green, 1998; Gordon y col., 1998), y segundo para compararla con las bases de datos disponibles (*GeneBank*) y comprobar si se trataba de una secuencia nueva o de un marcador de la zona ya descrito (algunos de los YAC utilizados para aislar marcadores, se sabía que contenían marcadores de la región como por ejemplo TGLA44). Para realizar esta comprobación también se utilizó el programa *FastA* del paquete GCG. En la Tabla IX se resumen los resultados de los microsatélites identificados.

A partir de la secuencia de cada clon, se diseñaron oligonucleótidos que flanqueaban la región del microsatélite, para poder analizar su polimorfismo vía PCR. En la Tabla X se muestran las secuencias de los oligonucleótidos diseñados para amplificar los microsatélites.

Los microsatélites identificados se nombraron como Bulge (*Bovine Université de Liège*) y el número de clon de donde proviene la secuencia.

En el apartado de Observaciones, de la Tabla IX, se indica que los clones 5 y 1, resultaron ser la misma secuencia, por ese motivo se nombrará a partir de ahora a los dos como Bulge05. Además en la secuencia del clon 27, se encontraron dos repeticiones tipo microsatélite, (CT)<sub>10</sub> y (GAG)<sub>8</sub>. La repetición (CT)<sub>10</sub> al tener mayor número de repeticiones fue elegida para utilizarla como marcador.

- Tabla VIII. Características de los clones positivos identificados.

Clon	STS	Microsatélite	Tamaño	levadura/no
1	Col5A2	(CT) <sub>12</sub>	160pb	no-levadura
5	Col5A2	(CT) <sub>12</sub>	90pb	no-levadura
18	TGLA44-BM81124	(GT) <sub>12</sub>	240pb	no-levadura
20 a	Col5A2	(CT) <sub>9</sub>	70pb	levadura
20 b	Col5A2	(CA) <sub>18</sub>	220pb	no-levadura
21	---	---	400pb	---
23	TGLA44-BM81124	(CA) <sub>13</sub>	500pb	no-levadura
24	---	---	500pb	---
25	---	---	350pb	---
26	TGLA44-BM81124	(CA) <sub>40</sub>	160pb	levadura
27	TGLA44-BM81124	(CT) <sub>10</sub>	160pb	no-levadura

- Tabla IX. Microsatélites identificados.

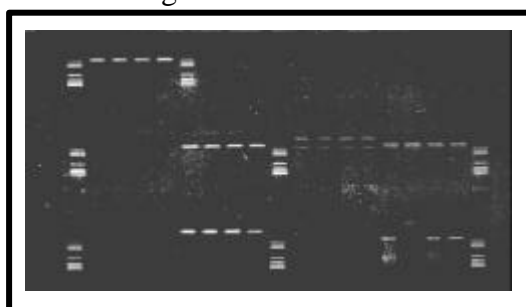
Clon	Microsatélite	Repetición	Observaciones
Clon1	Bulge05	(CT) <sub>12</sub>	Clon5=Clon1
Clon5	Bulge05	(CT) <sub>12</sub>	Clon5 = Clon1
Clon18	Bulge18	(GT) <sub>12</sub>	
Clon20b	Bulge20	(CA) <sub>18</sub>	
Clon23	Bulge23	(CA) <sub>13</sub>	
Clon27	Bulge27	(CT) <sub>10</sub>	(GAG) <sub>8</sub> en el mismo inserto
Clon28	Bulge28	(GT) <sub>15</sub>	

- Tabla X. Secuencias de los oligonucleótidos de los microsatélites identificados.

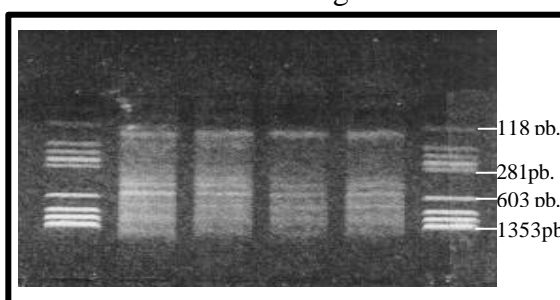
Clon5:	5up: 5'-GAT CAG GGA GTG GGA GAT CAG-3' 5dn: 5'- GTC ATA GGC TTG GCC CCA GTC CTT G-3'
Clon 18:	18up: 5'-ACA GTG AGT ACT TCT GTG AGC ACT C-3' 18dn: 5'-TGT GTT CAG AGA TAG CAT GTT GTA C-3'
Clon20:	20up: 5'-CAG CAG GTC TGT TGA AGT GTA TCA G-3' 20dn: 5'-AGT GGT AGC ATT CAC AGG TAG CCA G-3'
Clon23:	23up: 5'-ACA TTC TCT CAC CAA TAT GAC ATA C-3' 23dn: 5'-TAA GTC ACC ATT ACA TCC TTA GAA C-3'
Clon 28:	28up: 5'-AGG CAT ACA TCT GGA GAG AAA CAT G-3' 28dn: 5'-CAG AGG AGC CTG GCA GGC TAC CGT C-3'
Clon27:	27up: 5'-CTA CCT AAC AGA ATG ATT TTG TAA G-3' 27dn: 5'-AGT GTT CTT GCC TAG AGA ATC CCA G-3'

A continuación se probaron los oligonucleótidos diseñados en una PCR en condiciones estándar (Figura 20).

- Figura 20. Gel de agarosa con las PCR de algunos de los microsatélites.



- Figura 21. Amplificación del marcador Bulge05



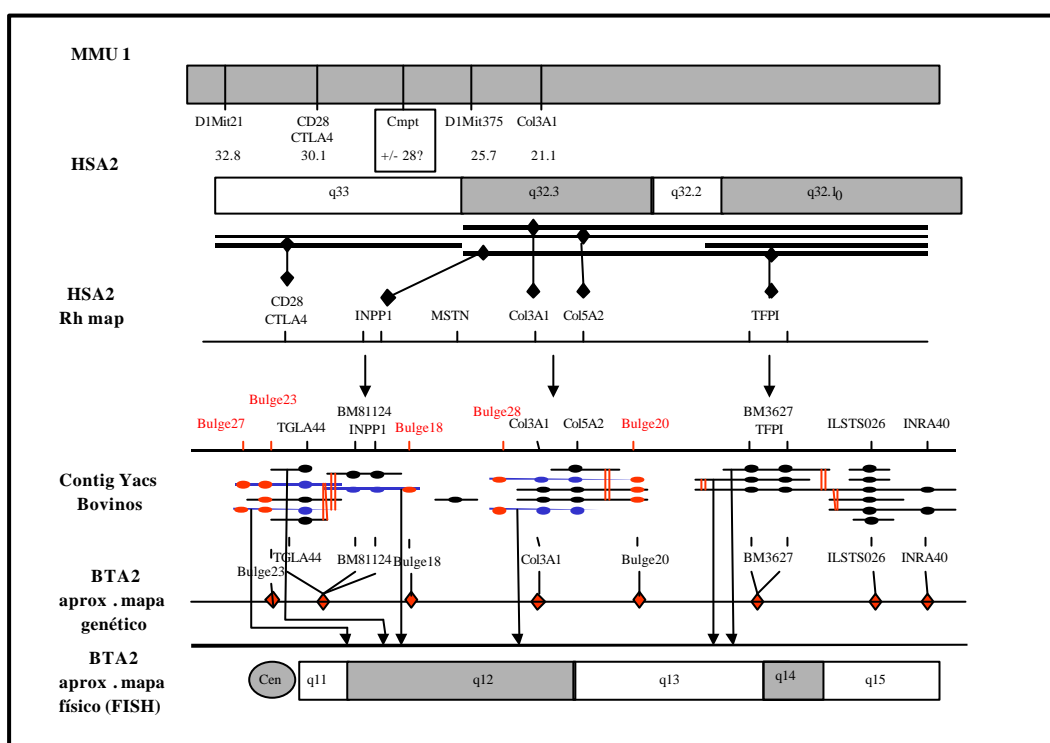
El clon 5 dio problemas en la amplificación, presentando un patrón de numerosas bandas (Figura 21). Se explicó este problema por tratarse de un marcador situado en el extremo de un elemento repetitivo. En bovino se ha estimado que el 45% de microsatélites están asociados a elementos repetitivos tipo SINE (Kaukinen y Varvio 1992, Vaiman y col. 1994, Kostia y col. 1997) y algunos tan cerca que uno de los oligonucleótidos se localiza dentro de uno de ellos, resultando amplificaciones inespecíficas. Por este motivo no volvió a ser utilizado (Figura 21).

### 3 Resultados.

Se han aislado 5 marcadores tipo microsatélite de la región BTA2q11-12. Estos marcadores se utilizaron en primer lugar en la construcción del *contig* de la región como STS en la genoteca de YAC. De los 5 microsatélites aislados, tres de ellos (Bulge18, Bulge20, Bulge23) resultaron polimórficos en las familias donde se utilizaron (familia de Sart Tilman, Charlier y col., 1995; familias IBRP, *International Bovine Reference Panel*, familias Maine-Anjou; Pirottin y col., 1999). Los estudios de ligamiento con estos marcadores en estas familias permitían localizar el *locus mh* entre los marcadores Bulge18 y Bulge20 en un intervalo de 3cM. (Pirottin y col., 1999).

La cartografía de la región cromosómica que contiene el *locus mh*, después de la adición de los nuevos marcadores, queda reflejada en la Figura 22.

- Figura 22. Cartografía de la región con los nuevos marcadores.



● YAC positivo a un STS.

|| YAC positivo por hibridación productos de PCR inter-SINE (Pirottin y col., 1999).

## Capítulo 2.

### Secuenciación del gen de la miostatina bovina y regiones flanqueantes.

---

#### 1 Antecedentes.

La hipertrofia muscular es debida a la presencia de mutaciones en la secuencia codificante del gen de la miostatina (Grobet y col., 1997; Kambadur y col., 1997; McPherron y col., 1997). El modo de herencia de la hipertrofia muscular es parcialmente recesivo, es decir para obtener un animal culón, éste debe heredar un alelo culón de cada uno de sus padres, lo que hace imposible obtener con toda seguridad un ternero culón en primera generación, si no es a partir de dos individuos culones. Sin embargo, el interés de proyectos encaminados a poder modificar este modelo de herencia para poder obtener terneros culones en primera generación a partir de hembras, por ejemplo de ganado lechero, de fenotipo normal, obliga a seguir adelante en el conocimiento de este gen. Por otra parte, la expresión del carácter culón también tiene aspectos negativos destacando los problemas de partos.

Existen individuos fenotípicamente hipermusculados en cuya secuencia de ADN codificante no se han encontrado mutaciones que expliquen ese fenotipo. Esto puede indicar que el fenotipo en estos animales esta causado por mutaciones en este mismo gen pero fuera de la zona codificante, o bien que se trate de fenocopias, es decir el mismo fenotipo pero con distinto determinismo genético (Grobet y col., 1998).

Para resolver estas dudas, era necesario determinar la secuencia completa del gen de la miostatina bovina y de sus regiones flanqueantes. Por eso, después de haber identificado el gen responsable y puestas en evidencia diferentes mutaciones responsables del fenotipo culón, se inició la secuenciación del gen y de sus regiones flanqueantes.

En dos de los artículos previamente citados (Grobet y col., 1997; Grobet y col., 1998), se utilizaron secuencias de la miostatina bovina. En el primero de ellos se obtuvo la secuencia completa del ADN codificante. En el segundo, con el fin de secuenciar completamente la región codificante de la miostatina a partir de ADN genómico, se localizaron las uniones exón/intrón y se demostró que el gen estaba compuesto por tres exones separados por dos intrones.

Existen secuencias de la miostatina bovina en *GenBank* con los siguientes números de acceso: AF019761 (Kambadur y col., 1997), AF019620 (McPherron y col., 1997).

Una vez localizadas las uniones exón/intrón, aprovechando la conservación de secuencias alrededor de estas uniones, se diseñaron oligonucleótidos para poder amplificar los dos intrones de la miostatina bovina a partir del YAC positivo a la miostatina que se encontraba en el *contig* de la región (Y189A3, Pirotin y col., 1999). Estos fragmentos de PCR se secuenciaron por el método de *primer-walking* siguiendo el protocolo de secuenciación *Dye terminator Cycle Sequencing Ready Reaction* (Perkin Elmer) utilizando un secuenciador automático ABI373.

### 1.1 Tamaño de la secuencia.

El objetivo era determinar la secuencia completa del gen de la miostatina y de sus regiones flanqueantes. Un aspecto importante era conocer el tamaño de las regiones flanqueantes al gen que había que secuenciar para tener una buena probabilidad de englobar al menos el gen completo y algunas de las posibles regiones reguladoras del gen de la miostatina. Se trataba de decidir el tamaño del clon que serviría de molde para la secuenciación, y por tanto el tipo de genoteca donde se debía buscar ese clon.

El promotor se puede describir como la región mínima con la que la enzima ARN polimerasa puede iniciar la transcripción, y está situado inmediatamente anterior (en 5') al inicio de la transcripción. Es decir normalmente secuenciando una o dos kilobases en 5' del inicio de la traducción deberíamos tener la región 5' no traducida con el inicio de la transcripción, y probablemente el promotor. En la región posterior o 3' la situación suele ser la misma. Después del codón Stop se necesitaba asegurar la existencia de una secuencia de varias kilobases también, para poder caracterizar la región 3' no traducida, y con ello, el lugar o lugares de poliadenilación que indican el final del transcrito. Además en la región en 3' de los genes también pueden asentarse regiones reguladoras de la transcripción (Wasserman y Fickett, 1998).

Se preveía que el tamaño del gen de la miostatina fuese menor de 7 kb.: tres exones de 500 pb. de tamaño medio (tamaño que se suponía en un principio sin contar las regiones 5' y 3' no traducidas), y dos intrones de 2 kb. Por lo tanto, la secuenciación de un clon de una genoteca de ADN genómico en fago  $\lambda$ , cuyo tamaño medio de inserto es 15 kb., debería satisfacer las exigencias que se habían propuesto. Por otro lado, una secuencia de ese tamaño, es suficiente para ser utilizada en posteriores protocolos de modificación del gen mediante estrategias de recombinación homóloga. Las estrategias de recombinación homóloga eran en principio las elegidas para posteriores protocolos de modificación del gen de la miostatina (ver por ejemplo Fernández y col., 2002). No se debe olvidar que el principal motivo que lleva a secuenciar completamente el gen de la miostatina es acumular el mayor conocimiento posible sobre las características de la secuencia, imprescindible a la hora de poder iniciar diferentes protocolos de modificación génica.

### 1.2 Estrategia de secuenciación.

Los actuales métodos de secuenciación permiten secuenciar como máximo cerca de una kilobase de una vez. Para secuenciar un fragmento de ADN muy grande (a partir de varias decenas de kilobases), existen básicamente dos estrategias: secuenciación directa de fragmentos de ADN por *primer-walking*, o secuenciación al azar por *Shotgun*.

- a) Secuenciación por *primer-walking*: Esta estrategia consiste en que una vez obtenidas una serie de secuencias iniciales, se siguen sucesivos pasos de secuenciación usando oligonucleótidos específicos diseñados a partir de la secuencia obtenida (Strauss y col., 1986). Las nuevas secuencias sirven para diseñar oligonucleótidos nuevos y continuar el proceso hasta obtener la secuencia completa.
- b) Secuenciación por *Shotgun* (Bankier y col. 1987; Andersson y col., 1997): en este caso se trata de obtener la secuencia completa a partir de las secuencias de fragmentos solapantes elegidos al azar, es decir dividir la secuencia total en una serie de secuencias parciales que solapen y más tarde unir las todas en una única secuencia final. El fragmento a secuenciar es cortado al azar por sonicación, se reparan los extremos de los fragmentos y se clonan en vectores pequeños para luego secuenciarlos (McMurray y col., 1998). Los mayores problemas en esta técnica residen en una ineficiente reparación de los fragmentos sonicados y en la aparición a menudo de poli-insertos en los vectores de clonado. Los datos obtenidos de las secuencias parciales solapantes deben ser alineados informáticamente para obtener la secuencia.

Los últimos proyectos de secuenciación del genoma de varias bacterias se han llevado a cabo mediante una estrategia de secuenciación al azar o *Shotgun* (Kaczorowski y Szybalski 1998; Fleischmann y col., 1995; Fraser y col., 1995; Bult y col., 1996).

Una vez revisados los métodos posibles de secuenciación, la secuenciación por *shot-gun* sería la de elección en estos casos. Sin embargo una experiencia anterior, para la secuenciación de un clon en fago  $\lambda$  del gen de la miostatina de ratón, nos hace pensar en desestimar una estrategia pura de *shot-gun sequencing*. En este caso particular, el ADN extraído del fago se fragmentó al azar por sonicación, los extremos de los fragmentos generados se repararon con la enzima *Klenow* para poder ser clonados, se seleccionaron por tamaño y se llevó a cabo un clonado en romo en el vector PUC18. Se eligieron un número de clones al azar, que representaban varias veces el tamaño de la secuencia final, y se alinearon todas las secuencias con los programas *Phred/Phrap/Consed package* (Ewing y col., 1998; Ewing y Green, 1998; Gordon y col., 1998). En este protocolo apareció el problema de la existencia de polinsertos (seguramente debido al clonado en romo). La aparición de polinsertos hace imposible la posterior alineación correcta de las secuencias (L. Grobet, comunicación personal). Por lo tanto en el caso de la secuencia de la miostatina bovina se evitó el clonado en romo. También existen referencias de aparición de problemas en la reparación de los extremos generados en un protocolo de sonicación (Nehls y Boehm, 1995).

## 2 Secuenciación del gen de la miostatina bovina y regiones flanqueantes.

Para evitar posibles errores en la secuencia del gen de la miostatina, se decidió no secuenciar a partir del YAC Y189A3 (Grobet y col., 1997; Pirottin y col., 1999) debido a los problemas de quimerismos descritos en los YAC, y si hacerlo a partir de clones en vectores que no presentasen estos problemas. Debido al tamaño previsto del



gen de la miostatina (ver página 52) se decidió secuenciar a partir de clones de genotecas comerciales de ADN genómico en vectores de reemplazamiento en fago  $\lambda$ . La estrategia elegida para secuenciar consistió en generar múltiples fragmentos al azar, por digestión con enzimas de restricción, que se secuencian siguiendo un protocolo de *primer-walking*. Posteriormente se obtiene la secuencia completa ensamblando las secuencias de estos fragmentos.

## 2.1 Obtención de un clon de fago $\lambda$ de la miostatina bovina.

Después de identificar un gen se deben obtener clones de ADN codificante o genómico a partir del análisis de genotecas. El tamaño pequeño del gen de la miostatina, posibilita que, un clon de una genoteca de ADN genómico en fago  $\lambda$  cuyo tamaño medio es de entre 15 y 20 kb., contenga el gen y una secuencia suficiente en ambos extremos, 5' y 3'. El tamaño también es compatible con un protocolo de recombinación homóloga, que puede utilizarse en estrategias de modificación del gen.

### 2.1.1 Genotecas bovinas de ADN genómico en fago $\lambda$ .

El genoma del bacteriofago  $\lambda$  es una molécula de ADN de doble cadena de aproximadamente 50 kb. de longitud. El ADN en las partículas víricas está contenido en forma de una molécula linear de doble cadena, con unos extremos cohesivos de cadena simple de 12 nucleótidos (Sambrook y col, 1989). Después de haber entrado en la célula hospedadora, los extremos cohesivos se asocian por complementariedad de bases, y forman una molécula circular de doble cadena. Las muescas en la molécula son rápidamente selladas por la ligasa de la bacteria hospedadora, para formar una molécula circular cerrada, que servirá de molde para la transcripción durante la fase temprana de infección. Durante esta fase se elige uno de los dos siguientes caminos:

- 1- Ciclo lítico: el ADN viral se replica muchas veces, se sintetiza una gran cantidad de productos víricos, se ensamblan las nuevas partículas víricas, y la célula se lisa, dejando libres una gran cantidad de nuevas partículas infecciosas.
- 2- Ciclo lisogénico: el genoma vírico infectivo se integra en el ADN bacteriano hospedador por unos lugares específicos de recombinación. Por lo tanto es replicado y transmitido a la progenie bacteriana como cualquier otro gen de la bacteria hospedadora. Durante esta fase sólo un gen del bacteriofago es expresado: el gen *cI*, que codifica para un represor que mantiene el ciclo lisogénico.

- Aspectos de la vida del fago a tener en cuenta para su cultivo.

- 1- Adsorción: El fago  $\lambda$  se adsorbe a los receptores de membrana externos de la bacteria hospedadora (*E.coli*) que normalmente se utilizan para el transporte de maltosa a la célula. La síntesis de estos receptores está inducida por la presencia de maltosa e inhibida por la glucosa, por eso los cultivos de bacterias hospedadoras deben hacerse en medios que contengan maltosa. Además la adsorción de las partículas de fago a los receptores se ve facilitada por los iones de magnesio, y se realiza eficaz y rápidamente tanto a temperatura ambiente como a 37° C. Sin

embargo la penetración de la célula y el transcurso del ciclo lítico no se realizan eficazmente a temperatura ambiente.

- 2- Lisis: la lisis de la célula hospedadora depende de dos proteínas víricas R y S. Mutaciones en el gen S previenen o retrasan la lisis, entonces permite continuar por un cierto periodo de tiempo el ensamblado de partículas víricas, sin que se lise la membrana celular, y se acumulan más partículas por bacteria hospedadora. Debido a este aumento en el rendimiento en número de partículas, la mayoría de vectores llevan la mutación S<sup>-</sup>.
- 3- Tamaño del inserto: La viabilidad del bacteriofago disminuye en gran medida cuando la longitud de su genoma es mayor de 105% o menor del 78% del tamaño del fago salvaje. Teniendo en cuenta que solo el 60% del genoma viral es absolutamente necesario para el ciclo lítico del fago (unas 20 kb. del brazo izquierdo, y entre 8-10 kb. del brazo derecho), podemos reemplazar la zona central del genoma por ADN externo. Así diferenciamos dos tipos de vectores tipo fago: vectores de inserción (por ejemplo  $\lambda$ gt10y  $\lambda$ gt22), que permite insertos de hasta 8 kb. y se suelen usar para elaborar genotecas de cDNA. El segundo tipo son los vectores de reemplazamiento (por ejemplo EMBL), donde al sustituir un parte del genoma del fago podemos utilizar insertos de mayor tamaño (15-20 kb.). Estos son los que se suelen usar para la fabricación de genotecas de ADN genómico.

- Características de las genotecas comerciales utilizadas

Genoteca Clontech BL1015j (EMBL3 SP6/T7).

Título:  $=10^8$  pfu/ml.

Vector de clonado: EMBL3 SP6/T7.

Diana de clonado: BamHI

Bacteria hospedadora: K802

Fuente de ADN: hígado de macho adulto

Número de clones independientes:  $2,3 \times 10^6$

Tamaño medio del inserto: 15 kb.

Tamaño de los insertos: 8-22 kb.

Genoteca amplificada una vez en la cepa K802.

Genoteca Stratagene 946702 (Lambda FIX II).

Título estimado:  $1,5 \times 10^{10}$  pfu./ml.

Vector de clonado: Lambda FIX II

Bacteria hospedadora: XL1-Blue MRA (P2)

Nº clones independientes:  $2,0 \times 10^6$  pfu.

Tamaño de los insertos: 9-23 kb.

Genoteca amplificada una vez en XL1-Blue MRA (P2)

### 2.1.2 Siembra de una genoteca de fagos.

Las genotecas comerciales de fagos se componen de un gran número de éstos diluidos en unos cuantos  $\mu$ l. de un tampón que los conserva, y un vial de la cepa bacteriana hospedadora. El número de clones independientes de que consta, el tamaño de estos clones y la cepa hospedadora están especificados. Normalmente están amplificadas, al menos una vez, en la cepa hospedadora. Esto implica que puede estar

representada de diferente manera una zona u otra del genoma. Además, suele acompañar la información del título aproximado de la genoteca. Sin embargo, es necesario hacer una dilución o diluciones seriadas y titularlas. A partir del título de esta dilución se procede a sembrar la genoteca.

Para conocer cuántos clones hay que sembrar, se suelen hacer los siguientes cálculos. En primer lugar hay que saber el tipo de genoteca, para saber la cantidad total que hay que representar. En el caso de genotecas de ADN genómico se sabe que el genoma de un mamífero tiene del orden de  $3 \times 10^9$  pb. Además hay que saber el tamaño medio de los insertos, y luego tener en cuenta que hay que sembrar varios equivalentes de lo que representa el genoma bajo estudio, normalmente 5. Esta redundancia se debe aumentar en el caso de que sean genotecas amplificadas. También hay que tener en cuenta el volumen de trabajo que se puede llevar a cabo en paralelo.

### 2.1.2.1 Titulación de la genoteca.

- En una placa de LB-agar, sembrar la cepa hospedadora. Mantener esta placa a 4°C.
- Picar una colonia y sembrar 4 tubos (Falcon 2059) con 5 ml. de medio líquido LB+maltosa 0,2% (para que se expresen sus receptores de membrana) +  $\text{MgSO}_4$  10mM. (que facilita la penetración del fago), y un control negativo.
- Incubar a 37°C. con agitación, 3 o 4 h., para coger al cultivo en fase de crecimiento.
- Diluir la genoteca bovina:
  - a) Dilución  $1/500 = 2 \mu\text{l. genoteca} + 1\text{ml } 1 \times \text{I dilution buffer}$ .  
Guardar a 4°C.
  - b) Dilución  $1/250.000 = 4 \mu\text{l. dilución a } (1/500) + 2\text{ml. } 1 \times \text{I dilution buffer}$ .  
Guardar a 4°C.
- Preparar 4 tubos (falcon 2059)

- Tabla XI. Titulación de la genoteca.

Nº tubo	1x I dilution buffer	cultivo bacteriano	dilución 1/250.000
1	100 $\mu\text{l.}$	200 $\mu\text{l.}$	2 $\mu\text{l.}$
2	100 $\mu\text{l.}$	200 $\mu\text{l.}$	5 $\mu\text{l.}$
3	100 $\mu\text{l.}$	200 $\mu\text{l.}$	10 $\mu\text{l.}$
4	100 $\mu\text{l.}$	200 $\mu\text{l.}$	---

- Incubar los tubos al baño maría 37°C., 15'. Durante este tiempo los fagos se adsorben a la membrana de la célula hospedadora mediante los receptores externos del transporte de la maltosa.
- Añadir 3 ml. LBsoft top agar a 50°C. y mezclar. La temperatura no debe ser mayor de 50°C. para evitar matar a las células.
- Sembrar la mezcla en placas de LB-agar+ $\text{MgSO}_4$ .

En las placas del 1 al 3 se cuentan las placas de lisis. La placa nº 4 sirve de control de la temperatura. En ella se debe encontrar un tapiz bacteriano completo en

toda la placa, si no es así es porque se está trabajando con una temperatura demasiado alta.

5ª placa con 3ml. LB-soft Top agar. Esta placa sirve para diferenciar una placa sólo con LB-soft Top agar, de una con tapiz bacteriano, y poder controlar bien lo que ha ocurrido en las placas anteriores.

6ª placa, sin nada, control (-).

- Incubar a temperatura ambiente 10' hasta que se enfríe el medio.  
Dar la vuelta a las placas e incubar a 37°C. toda la noche.

- Contado de las placas de lisis por cada dilución y estimación del número de pfu. por unidad de volumen de la genoteca. Con estos datos se hace una dilución de trabajo de alrededor de 10.000 pfu./µl.

### 2.1.2.2 Siembra de la genoteca.

- Cálculo del número de clones a analizar:

$3 \times 10^9$  pb. (tamaño genoma de mamíferos)  $\times$  7 (la genoteca está amplificada una vez)/15 Kb. (tamaño medio de los clones) = 1,4 millones es decir, sembrar entre 1,5 –2 millones de clones.

Se preparan 30 placas de LB-agar+MgSO<sub>4</sub>. Se hace un cultivo líquido de la cepa hospedadora en LB+maltosa+MgSO<sub>4</sub>, a partir de una colonia individual. Incubar a 37°C. con agitación 5h.

Autoclavar una botella de 1 litro vacía, donde se llevará a cabo la incubación de los fagos con la bacteria hospedadora. Se preparan 700 ml. de LB-soft top agar + MgSO<sub>4</sub> 10 mM., se autoclava también y se mantienen a 50°C. al baño maría.

- Preparar la siembra:

se van a sembrar 1,5 millones de clones / 75.000 clones/placa = 20 placas (22 x 20 cm.)

En una botella vacía autoclavada se mezclan las bacterias hospedadoras crecidas en presencia de maltosa y MgSO<sub>4</sub> y los fagos. Se incuba 15' a 37°C., en este momento se produce la adsorción de los fagos a los receptores de la maltosa de la bacteria.

Luego se añade el medio semilíquido (LB-soft top agar+ MgSO<sub>4</sub>) caliente para poder sembrar y distribuir la mezcla homogéneamente en las placas. Se mantiene 10' a temperatura ambiente, y una vez solidificado se da la vuelta y se incuba a 37° C. toda la noche. Durante la incubación los fagos adsorbidos penetran en la bacteria y empiezan en ella el ciclo lítico. Una vez acabado, son liberados e infectan las células vecinas. Por eso cada clon individual está representado por lo que se llama placa de lisis o pfu., unidades formadoras de placas, que no es más que un hueco en el tapiz bacteriano provocado por la lisis de bacterias, donde hay muchos fagos concentrados que proceden del mismo fago inicial, luego clones del primero que se adsorbió en la primera incubación en medio líquido.

Las placas de la genoteca sembrada se guarda a 4°C.

### 2.1.2.3 Conservación de la genoteca.

En la siembra de la genoteca no se emplea ningún tipo de antibiótico. Después de cierto tiempo pueden aparecer las primeras contaminaciones en las placas, sobre todo debido a la presencia de hongos. Para conservar las placas se hace lo siguiente:

Se recorta un papel 3MM de un tamaño algo menor que las placas y se empapa en cloroformo. Se coloca cada placa de la genoteca boca-abajo sobre el papel empapado en cloroformo, y se deja actuar los vapores durante 5'. Luego se retira la placa y se le coloca la tapadera.

**2.1.3 Réplicas de placas de genoteca de fago.** (membrana de nailon Hibond-N, Amersham LIFE SCIENCE).

a) Membranas primarias.

- Las placas deben permanecer al menos 2 h. a 4° C. antes de empezar a hacer las réplicas.
- Identificar la membrana.
- Depositarla sobre la placa, y dejarla durante 5'.
- Hacer marcas con una aguja de forma no simétrica con el fin de poder situarla después para identificar los positivos.
- 5' sobre papel 3MM empapado en solución desnaturalizante.
- 5' sobre papel 3MM empapado en solución neutralizante.
- Lavado en SSC 2x.
- Dejar secar a temperatura ambiente sobre un papel 3MM seco.
- Formar bocadillos con papel 3MM y las membranas.
- Fijar el ADN a las membranas incubando a 80° C. 2 h.

### 2.1.4 Aislamiento de clones positivos.

- Hibridación de las membranas

Como sonda de la miostatina bovina se utilizaron dos fragmentos solapantes de ADN codificante de la miostatina, amplificados utilizando los oligonucleótidos GDF8.27-GDF8.12 (753 pb.) y GDF8.11-GDF8.21 (724 pb.) (Grobet y col., 1997), y marcados siguiendo el protocolo de marcado por oligos (*oligolabeling*, Random Primers DNA Labeling System, GIBCOBRL, Life Technologies; ver página 42).

Tras la primera hibridación, se identifican las placas de lisis positivas, pero normalmente es imposible diferenciar el clon individual positivo, ya que se trata de una siembra a alta densidad. Esto permite en unas pocas placas (unas 20) poder sembrar toda la genoteca, pero tiene el inconveniente de obligar a rondas sucesivas de aislamiento, siembra e hibridación hasta poder aislar un clon individual.

Se extrae una zona pequeña de la placa donde se encuentra el clon positivo taladrando la superficie de agarosa con la parte posterior de una punta azul autoclavada, y se introduce en 1 ml. de 1x **I** *dilution buffer*. Se incuba a 37° C. con agitación ligera

durante 4 h. para permitir a los fagos que salgan de la agarosa al medio tamponado. Esta solución conteniendo los fagos se puede guardar a 4°C.

Los fagos aislados se siembran en una placa a menor densidad, para poder picar esta vez la placa de lisis positiva aislada.

Se diluyen los positivos primarios y se siembran según la Tabla XII:

A) 1/500: 2 µl. + 1 ml 1x *I dilution buffer*

B) 1/50.000: 2 µl. dilución A + 200 µl. 1x *I dilution buffer*

- Tabla XII. Siembra secundaria en placas de Petri redondas grandes.

Nº tubo	1x <i>I dilution buffer</i>	bacterias	dilución B
*(-)	52 µl.	---	---
*Bact.	152 µl.	300 µl.	---
(+)	690 µl.	1520 µl.	70 µl.

\*Siembras en placas petri pequeñas.

- Hacer las replicas en membrana de la siembra secundaria.
- Hibridación de las membranas secundarias.

En este momento se pica una placa de lisis positiva aislada que contendrá solamente un clon de fagos positivos. Ahora, se taladra la agarosa con una punta amarilla cortada. Se introduce en 500 µl. de 1x *I dilution buffer* y se incuba a 37° C. con agitación durante 4 h. Se añaden 30 µl. de cloroformo, con una finalidad doble: por un lado matar las bacterias, y por otra precipitar el conjunto de proteínas en solución (cápsides virales mal formadas o no acabadas).

Se vuelven a diluir los clones individuales, para hacer una tercera comprobación. En esta siembra terciaria todas las placas de lisis deben ser positivas en la hibridación (Tabla XIII). Si no fuera así, se volvería a picar un clon individual positivo, se repetiría el proceso hasta tener una placa con todos las placas de lisis positivas.

Para diluir los clones individuales se tiene en cuenta lo siguiente:

- 1 placa de lisis contiene aproximadamente  $10^7$  fagos
- $10^7 / 500 = 2 \times 10^4$  fagos / µl.
- dilución (a): 50 µl. + 950 1x *I dilution buffer*  $10^3$  fagos / µl.
- dilución (b): 5 µl. de (a) + 495 1x *I dilution buffer* 10 fagos / µl.

- Tabla XIII. Siembra terciaria (placas petri redondas grandes).

Tubo	1x <i>I dilution buffer</i>	bacterias	dilución (b)
(+)	200 µl.	600 µl	100 µl
Bact.	300 µl.	600 µl.	---

- Hacer las membranas terciarias.
- Hibridación filtros terciarios.
- Picar varias placas de lisis positivas de cada clon en un tubo con 1ml. de 1x *I dilution buffer*, y añadir cloroformo.

## 2.2 Clones positivos de la miostatina bovina.

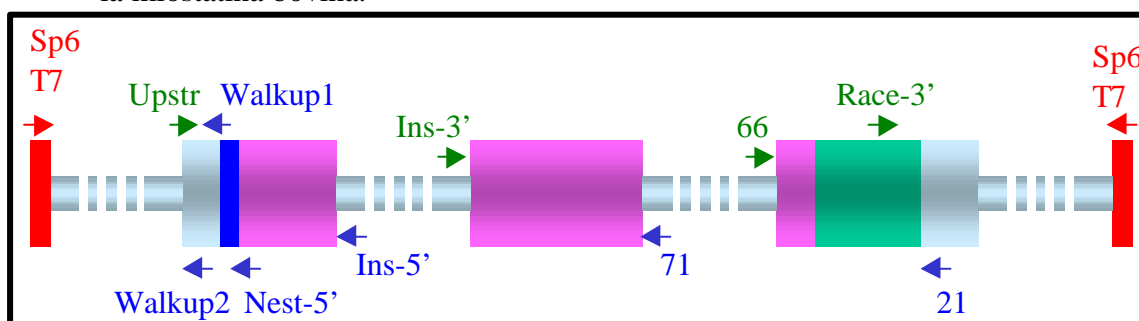
Se identificaron 5 clones positivos (Tabla XIV) que se caracterizaron por PCR utilizando los oligonucleótidos que se muestran en la Figura 23. Por caracterizar los clones se entiende, estimar el tamaño del clon completo, intentar conocer si contienen el gen completo, y el tamaño de la secuencia que contienen en los extremos 5' y 3'. La información obtenida por las PCR para caracterizar los clones, así como los protocolos de PCR utilizados se muestra en la Tabla XV

No se encontró ningún clon que tuviera el tamaño adecuado en ambos extremos. Para poder tener una secuencia suficientemente grande de los dos extremos, se tuvo que recurrir a seleccionar dos clones, uno para secuenciar el extremo 5', el clon **B-21**, y otro para secuenciar el extremo 3', en este caso el clon **A-13** (Figura 24). Una vez elegidos los clones, se amplificaron por PCR con varios de los oligonucleótidos situados en la secuencia y se eligieron los que amplificaban de una forma más específica. Estos fragmentos de PCR se convirtieron en el molde a partir del cual se obtuvo la secuencia de los extremos del gen. Las secuencias de intrones y exones ya estaban disponibles, y una vez obtenidas una secuencia de cada uno de los extremos del gen, se alinearon todas ellas para formar una única secuencia completa.

- Tabla XIV. Clones positivos de miostatina bovina.

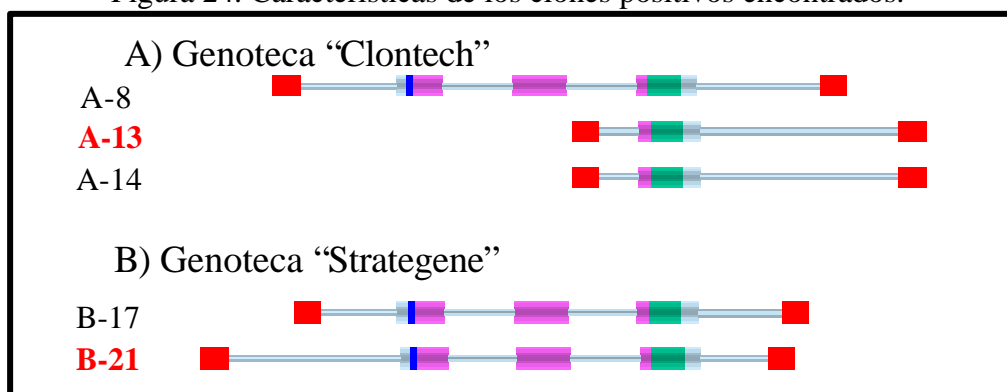
Genoteca	Nº Clones analizados	Clones positivos
A) Clontech	1,5 x 10 <sup>6</sup>	3: A-8, A-13, A-14
B) Stratagene	2 x 10 <sup>6</sup>	2: B-17, B-21

- Figura 23. Localización de los oligonucleótidos utilizados sobre la secuencia de la miostatina bovina.



- En rojo los oligos situados en los brazos del vector. En azul los oligos situados dentro de la secuencia en sentido 3'-5', y en verde los oligos situados dentro de la secuencia en sentido 5'-3'. Sp6 y T7 son oligonucleótidos situados en los brazos del vector de clonado, por eso podrían estar situados en cualquiera de los dos lados dependiendo de la orientación del inserto bovino.

- Figura 24. Características de los clones positivos encontrados.



- Tabla XV. Protocolos de PCR para caracterizar clones positivos.

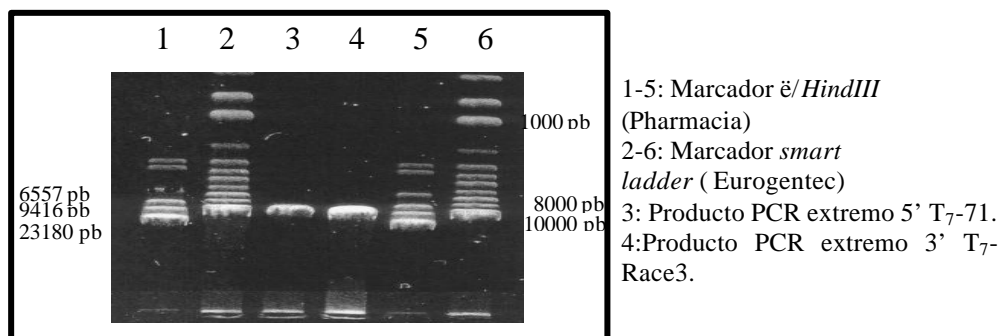
Protocolo de PCR	Indicaciones de un resultado (+)
T7-Sp6	Tamaño del clon completo
T7-Race3' Sp6-Race3' T7-Nest5' Sp6-Nest5'	Estas 4 PCR sirven para orientar el clon dentro del vector, y conocer el tamaño de la secuencia de los extremos 5' y 3'.
Upstr-Ins5'	El clon correspondiente contiene el exón 1
Ins3'-71	El clon correspondiente contiene el exón 2
66-21	El clon correspondiente contiene el exón 3

### 2.3 Amplificación de los clones elegidos .

Los productos PCR de cada extremo se secuenciaron utilizando una estrategia combinada de *Shot-gun* y *primer-walking*. Cada producto se digirió con enzimas de restricción para dividirlo en fragmentos más pequeños, que se secuenciaron. A partir de estas secuencias se diseñaron oligonucleótidos en los dos extremos de la secuencia obtenida. Esto permitió iniciar una estrategia de *primer-walking* a partir de varios puntos diferentes, utilizando como molde el producto PCR inicial. Más tarde, todas estas secuencias generadas a partir del producto PCR inicial de cada extremo, se alinearon utilizando una aplicación informática (programas *Phred/Phrap/Consed package*; Ewing y col., 1998; Ewing y Green, 1998; Gordon y col., 1998) hasta obtener un solo *contig* de secuencias de cada uno de los extremos del gen, es decir de cada uno de los productos PCR. Estas secuencias se utilizaron para obtener las secuencias completas de los extremos 5' y 3' del gen. El resto de la secuencia de la miostatina bovina ya se conoce (ADN codificante, Grobet y col. 1997; Kambadur y col., 1997; McPherron y col. 1997, Grobet y col., 1998, para los intrones ver más adelante) y se agregará posteriormente para formar una única secuencia consenso total de la miostatina bovina.

Una vez elegidos los clones para cada extremo, se llevó a cabo una PCR en condiciones de *Long-Range PCR*, utilizando un oligonucleótido dentro del vector (T7, en este caso, debido a la orientación del clon dentro del vector) y diferentes oligonucleótidos situados de forma consecutiva dentro de la secuencia conocida. De entre ellos se eligió el que produjera mejor resultado en la PCR

- Figura 25. Amplificación de los fragmentos de PCR elegidos, en condiciones de *Long Range PCR*.





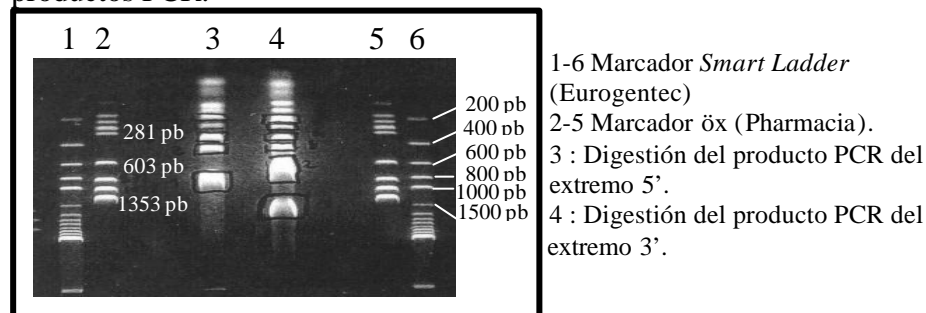
A la vista de los resultados, los protocolos de PCR elegidos para amplificar cada extremo del gen de la miostatina fueron:

- Extremo 3' del gen de la Miostatina: PCR del clon A-13 con los oligos T7 y Race3'.
- Extremo 5' del gen de la Miostatina: PCR del clon B-21 con los oligos T7 y 71.

## 2.4 Digestión de los productos PCR y clonado de los fragmentos de digestión.

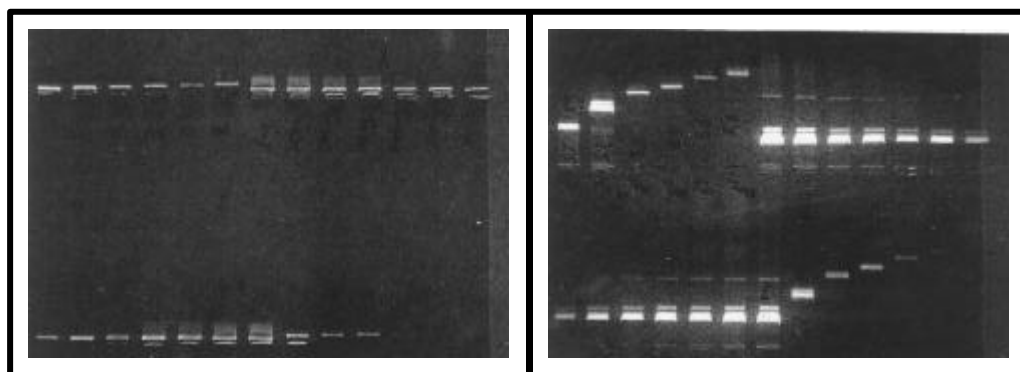
Los productos PCR se digirieron con varios enzimas de restricción: *EcoRI* (GibcoBRL), *BglIII* (GibcoBRL), *MboI* (GibcoBRL) y *BamHI* (GibcoBRL). Se eligió la digestión con *MboI* que generaba un patrón de entre 5 y 10 bandas. Por lo tanto se digirieron ambos productos *Long-Range PCR* con *MboI*.

- Figura 26. Digestión *MboI* de los productos PCR.



Tras la digestión con *MboI* de los dos fragmentos, se separaron en gel de agarosa las diferentes bandas obtenidas, se recortaron del gel y se extrajo el ADN. Cada uno de los segmentos del patrón de digestión se clonó separadamente en PUC18/*BamHI*-BAP, teniendo en cuenta que las dos enzimas *MboI* y *BamHI* generan extremos cohesivos.

- Figura 27. ADN de cada una de las bandas del patrón de la digestión.



- Arriba y de izquierda a derecha: fragmentos de digestión del extremo 3' preparado para clonar, y marcadores a concentraciones conocidas decrecientes.
- Abajo y de izquierda a derecha: marcadores a concentraciones conocidas decrecientes, y fragmentos de digestión del extremo 5' preparado para clonar.
- Gel de la izquierda, migración de 5 minutos. Gel de la derecha, migración de 1 hora.

Una vez clonados cada uno de los segmentos por separado, se picaron, al menos, 4 colonias individuales por banda del producto clonado (ver Figura 27).

En el extremo 5' se clonaron por separado 5 segmentos recortados del gel de agarosa (5'-1 a 5'-5). En el segmento 1 se pueden apreciar 2-3 bandas diferentes, por lo tanto se picaron 3 bandas x 4 colonias individuales por banda = 12 colonias. En los segmentos 2, 4 y 5 se distinguen 2 bandas, con lo cual se picaron 8 colonias individuales. Y en el segmento 3, una sola banda, luego 4 colonias. En el extremo 3' se clonaron 6 segmentos diferentes del gel por separado y se siguió el mismo criterio a la hora de cuantas colonias individuales picar. Se redondeó para acabar completando una microplaca de 96 pocillos (Tabla XVI).

- Tabla XVI. Microplaca con clones individuales elegidos.

- Número de colonias individuales para picar:

5'-1: 3 x 4 = 12	---- 15
5'-2: 2 x 4 = 8	---- 10
5'-3: 1 x 4 = 4	---- 5
5'-4: 2 x 4 = 8	---- 10
5'-5: 2 x 4 = 8	---- 10
3'-1: 1 x 4 = 4	---- 6
3'-2: 2 x 4 = 8	---- 10
3'-3: 2 x 4 = 8	---- 10
3'-4: 1 x 4 = 4	---- 5
3'-5: 1 x 4 = 4	---- 5
3'-6: 2 x 4 = 8	---- 9
Control (-) =	1
Total = 96 colonias (una microplaca).	

Los clones se sometieron a una amplificación por PCR para estimar el tamaño del inserto y al mismo tiempo poder elegir los que fueran diferentes, ya que en algunos casos los productos de clonado estaban constituidos por más de una banda. Se eligieron los siguientes clones para secuenciar (identificados por sus coordenadas en la microplaca, Tabla XVII):

- Tabla XVII. Clones elegidos para secuenciar

Clones 5'	Clones 3'
5.1: E1, H1, A2, C2, A1.	3.1: C7, H7.
5.2: D3, E3, H3, A4.	3.2: C8, E8, A9.
5.3: D4, F4.	3.3: D9, B10.
5.4: G4, H4, C5, G5.	3.4: E10, F10.
5.5: B6, C6, G6	3.5: C11, E11.
	3.6: B12, A12, C12.

A continuación se secuenciaron dichos clones con los oligos universales, M13-*forward* y M13-*reverse* localizados en los brazos del vector PUC-18, y se alinearon las secuencias con la ayuda del grupo de programas *Phred/Phrap/Consed package* (Ewing y col., 1998; Ewing y Green, 1998; Gordon y col., 1998). Algunas de las secuencias obtenidas se alinearon con otras empezándose a formar pequeños *contig* o fragmentos

de secuencias. Se obtuvieron los siguientes fragmentos, donde se indica las secuencias que lo forman y el tamaño (Tabla XVIII):

- Tabla XVIII. Fragmentos de secuencias a partir de clones.

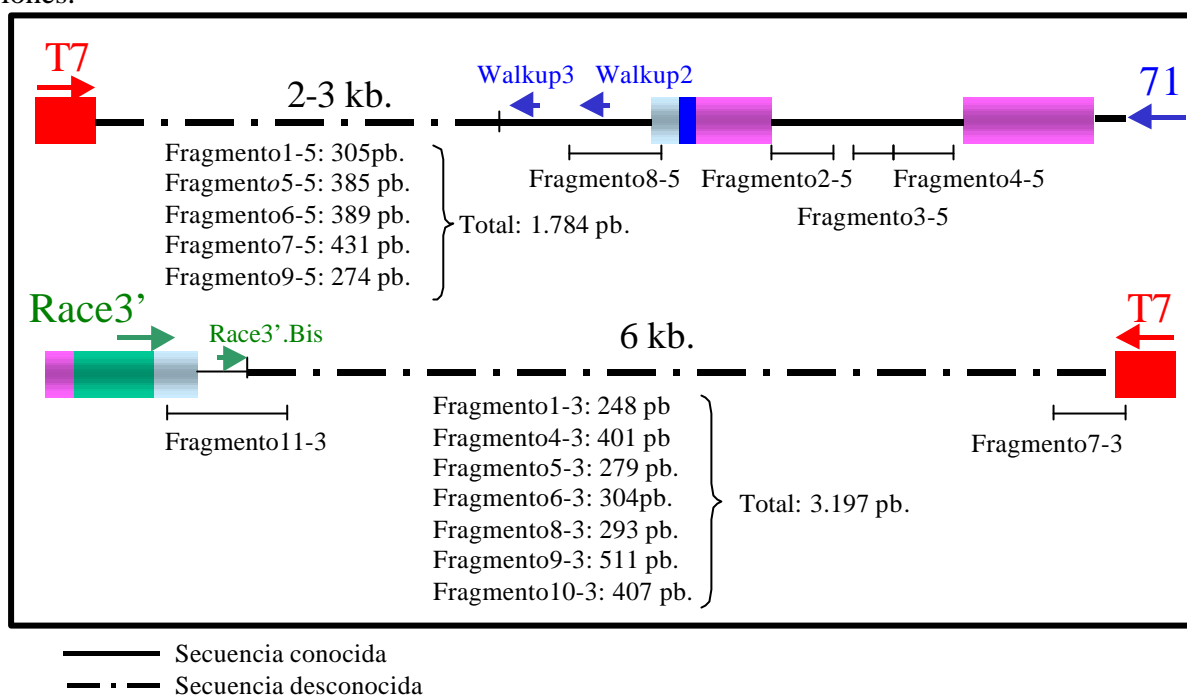
Extremo 5' (total secuenciado 3.915 pb.)	Tamaño
Fragmento1: A4-rev + A4-for	442 pb.
Fragmento2: H1-rev + A2-rev	335 pb.
Fragmento3: H1-for + A2-for	395 pb.
Fragmento4: A1-rev + A1-for	799 pb.
Fragmento5: C2-for + E1-for	385 pb.
Fragmento6: F4-rev + F4-for	389 pb.
Fragmento7: H3-for + D3-for + E3-for	431 pb.
Fragmento8: DraI-2 + H4-for + DraI-1 + G6-for	465 pb.
Fragmento9: C5-rev + G5-rev + D4-for + G4-for + G5-for + C5-for + D4-rev	
Extremo 3' (total secuenciado 3.197 pb.)	Tamaño
Fragmento1: B12-for	248 pb.
Fragmento2: A12-for (secuencia defectuosa)	
Fragmento3: A9-rev + E8-rev (secuencia defectuosa)	
Fragmento4: A9-for + E8-for	401 pb.
Fragmento5: C7-for + H7-rev	279 pb.
Fragmento6: C8-rev + C8-for	304 pb.
Fragmento7: C12-seq24 + C12-rev + 3T7	428 pb.
Fragmento8: E11-rev + E11-for + C11-for	293 pb.
Fragmento9: B10-rev + D9-seq24 + B10-seq24	511 pb.
Fragmento10: F10-for + F10-rev + E10-rev + E10-for	417 pb.
Fragmento11: 5up2.1 + 4up2.1 + 3up2.1 + 5race1 + 4race1 + 3race1 + C7-rev + H7-for + 3R3bis	686 pb.

Estos fragmentos de secuencias de clones se alinearon con las secuencias ya conocida de la miostatina (parte de los exones traducidos e intrones). Algunos de ellos se situaron en parte de la secuencia ya conocida, estas secuencias no se utilizaron. El resto de fragmentos se utilizaron para diseñar oligonucleótidos, que sirvieran para secuenciar los extremos 5' y 3' del gen de la miostatina en las dos hebras, a partir de los productos PCR de cada extremo. Estos oligonucleótidos se utilizarían para iniciar la secuenciación por *primer-walking* a partir de producto PCR exclusivamente.

La secuencia de la miostatina bovina resultante, por lo tanto, será una secuencia obtenida a partir de un producto PCR de un protocolo *Long-Range PCR* y no a partir de clones, para evitar problemas de falsos polimorfismos debidos al posible clonado de productos de PCR con mutaciones producidas por errores de la polimerasa.

En la Figura 28 se puede ver la localización de los fragmentos obtenidos sobre las secuencias conocidas de la miostatina bovina. En trazo continuo se representan las secuencias conocidas y en trazo discontinuo se representa la secuencia aún no conocida. Los análisis de homología de secuencias permitieron localizar algunos fragmentos en regiones conocidas de la miostatina bovina o en los brazos del vector. Sin embargo en otros casos no pudieron ser ubicados, éstos estarán localizados en algún lugar dentro de la zona representada con trazo discontinuo.

- Figura 28. Localización de los fragmentos obtenidos a partir de secuencias de los clones.



## 2.5 Secuenciación a partir de los productos PCR.

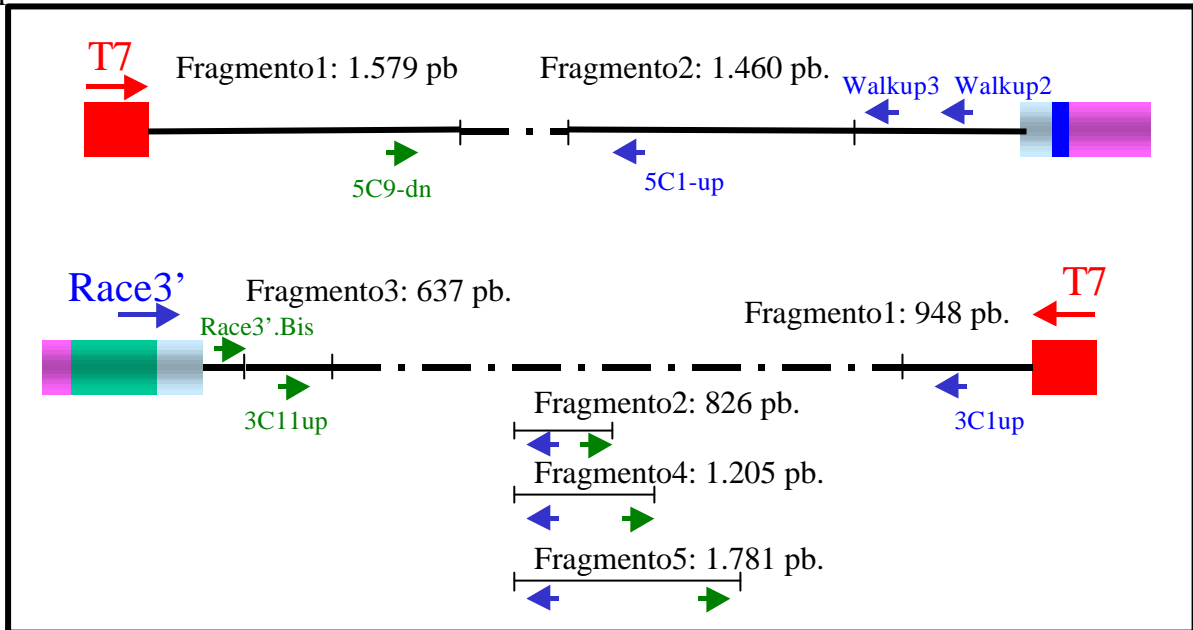
En los fragmentos obtenidos (Figura 28) se diseñaron oligonucleótidos de tal manera que siempre existiera una secuencia conocida que permitiera controlar el alineamiento posterior de las nuevas secuencias con las ya conocidas. Es decir, se colocaba siempre el oligonucleótido con unas 20-30 bases de secuencia de buena calidad delante de él, para estar seguro que la aplicación informática utilizada tenía margen para superponer la secuencia ya conocida con la siguiente. En los fragmentos de situación conocida unidos a secuencia conocida (Fragmentos 11-3 y 7-3) se diseñaron los oligonucleótidos para avanzar hacia la zona de secuencia desconocida. En los fragmentos de situación desconocida (Fragmentos 1-5, 5-5, 6-5, 7-5, 9-5, 1-3, 4-3, 5-3, 6-3, 8-3, 9-3, 10-3) se diseñaron oligonucleótidos en los dos extremos del fragmento. Otros fragmentos se situaban dentro de secuencias conocidas no se tuvieron en cuenta. A partir de ahora se secuenció con estos oligonucleótidos diseñados, utilizando como molde los productos PCR elegidos.

En la Figura 29 se puede ver el avance de la secuencia conocida después de la primera ronda de secuenciación. Cada uno de los fragmentos está compuesto por el brazo del vector en un extremo, y el límite de la secuencia conocida por otro. Además se localizan algunos de los oligonucleótidos ya utilizados sobre la secuencia. Del mismo modo que en la Figura 28 se representa la secuencia conocida con trazo continuo y la todavía desconocida con trazo discontinuo. Además se puede estimar la secuencia que falta por conocer, restando al tamaño estimado del fragmento por PCR la suma total de los fragmentos.

En el extremo 5' tras alinear todas las secuencias, se obtuvieron dos únicos fragmentos, cada uno de ellos unido a las secuencias conocidas de cada lado, uno a secuencias enlazadas con el brazo del vector (Fragmento1), y el otro enlazado a las

secuencias junto al exon1 (Fragmento2). Se diseñaron *primer-walking* en los extremos para avanzar hacia la secuencia desconocida (5C9dn / 5C1up).

- Figura 29. Avance de la secuencia tras la primera ronda de secuenciación a partir de producto PCR.



En el extremo 3' la situación tras esta primera ronda de secuenciación era diferente. Se tenían dos fragmentos orientados y situados en una parte y otra del fragmento:

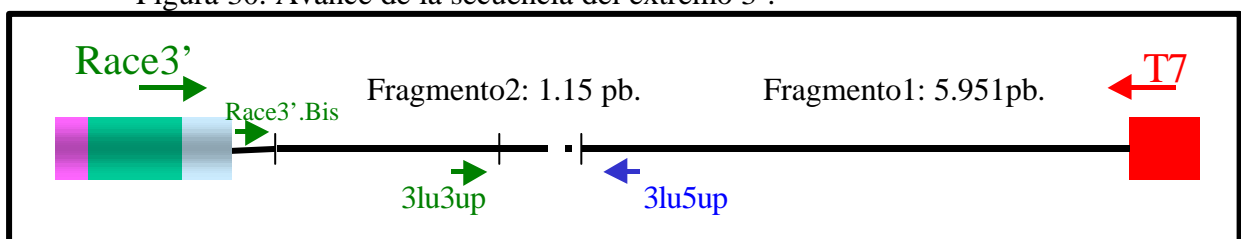
- Fragmento3 unido a las secuencias del exón 3 del gen,
- Fragmento1 con secuencias del extremo del vector de clonado.

Por otro lado, se tenía una serie de fragmentos de los que no se podía conocer ni su localización ni su orientación, debiéndose situar en algún lugar entre los dos fragmentos señalados anteriormente. Se diseñaron igualmente *primer-walking*:

- En el caso de los dos fragmentos de situación conocida, solo en un extremo hacia la secuencia desconocida,
- en los fragmentos de situación y orientación desconocida, en ambos lados.

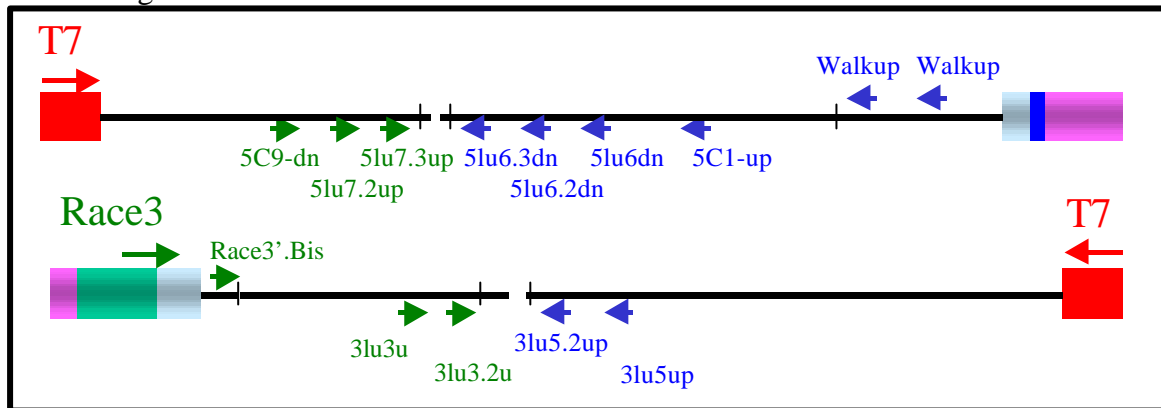
Tras esta segunda ronda de secuenciación, se localizaron y orientaron esos fragmentos de la región 3' del gen y se diseñaron *oligos-walking* para avanzar hacia la secuencia desconocida (3lu3up / 3lu5up).

- Figura 30. Avance de la secuencia del extremo 3'.



Tras varias rondas de *primer-walking* con los oligonucleótidos diseñados en los extremos de los fragmentos, la secuencia quedó como se ve en la Figura 31. Cada oligonucleótido en el gráfico supone una ronda de *walking*.

- Figura 31. Avance de la secuencia tras varias rondas de secuenciación.



Así se continuó hasta acabar formando un solo fragmento de cada uno de los extremos.

## 2.6 Ensamblado de todas las secuencias obtenidas.

Todas las secuencias obtenidas se analizaron con el grupo de programas *Phred/Phrap/Consed package* (Ewing y col., 1998; Ewing y Green, 1998; Gordon y col., 1998). Con el propósito de obtener un único fragmento de secuencias, se introdujeron secuencias obtenidas a partir de varios moldes diferentes:

- Secuencia del ADN codificante de la miostatina bovina de un individuo de fenotipo normal (Grobet y col, 1997).
- Secuencias de los exones de un gran número de individuos que fueron utilizados para detectar mutaciones (Grobet y col., 1998).
- Secuencias de los intrones secuenciados a partir de productos de PCR procedentes de YAC (Y189A3 que contiene el gen de la miostatina, Pirotin y col., 1999)
- Secuencias obtenidas con el protocolo de secuenciación de los extremos del gen antes descrito.

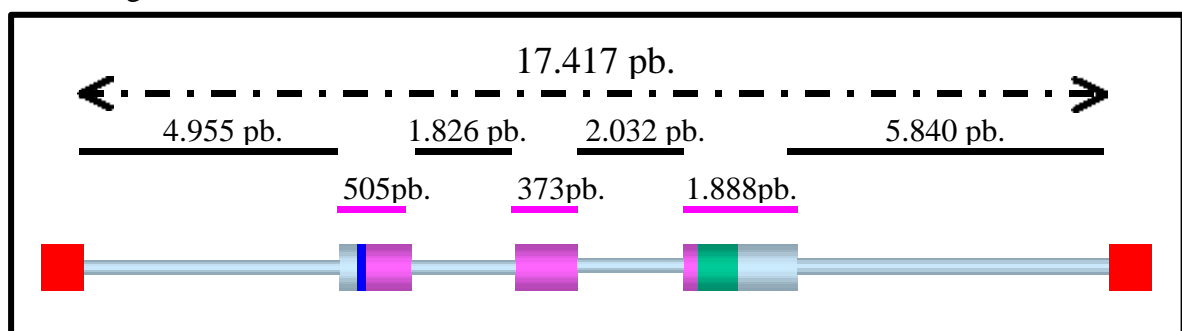
La secuencia consenso final estaba constituida, en su mayoría, por secuencias de las dos hebras y en algunos casos por más de dos secuencias complementarias.

En un primer análisis, con todas las secuencias introducidas aparecieron seis fragmentos que cubrían la secuencia completa. Se procedió a una fase de limpieza, donde se eliminaron del análisis todas las secuencias de baja calidad. Tras esta labor, la secuencia completa se cubrió con tres fragmentos, que el grupo de programas *Phred/Phrap/Consed package*, era incapaz de unir. Sin embargo, un análisis más detallado de los límites de estos fragmentos permitió unirlos en una única secuencia consenso:

- El primer fragmento estaba formado por secuencias desde el extremo del clon en fago lambda en 5' hasta una zona repetitiva, compuesta por 12 T seguidas, y el segundo empezaba precisamente en esa zona de poli-T. Estas zonas de polinucleótidos hacen que la secuenciación en ellas sea de una calidad baja, debido a un problema de la polimerasa para leer correctamente 12 T seguidas y hace que no todas las moléculas contengan el mismo número de T. La secuencia a partir de entonces se compone de una mezcla de moléculas y da una imagen de secuencias superpuestas, lo que se traduce en un gráfico de mala calidad que la aplicación informática no tiene en cuenta.
- La unión entre el fragmento 2 y el 3, era el resultado de algo similar. Ese tramo de secuencia por ambos lados era de una baja calidad no llegando al límite exigido por el programa, por tanto no las tiene en cuenta y no es capaz de unirlos. Sin embargo se tenían dos pruebas a favor de que esos dos fragmentos estaban verdaderamente unidos:
  - Los dos fragmentos se alineaban cuando se utilizaba otro programa que no tiene en cuenta la calidad de la secuencia, como es el programa *Bestfit* del grupo de programas GCG. Este programa busca el mejor segmento de alineamientos posible entre dos secuencias dadas. Los alineamientos se hacen insertando huecos para maximizar el número de coincidencias en un segmento. Las secuencias pueden ser de diferentes tamaños y tener solamente un segmento pequeño de similitud entre ellas. Los alineamientos se hacen utilizando el algoritmo *local homology* de Smith y Waterman (Smith y Waterman, 1981).
  - El tamaño del fragmento, teniendo en cuenta que el solapamiento del programa *Bestfit* sea correcto, coincide con el tamaño del producto PCR a partir de ADN genómico y de ADN del clon de fago a partir del cual se estaba secuenciando.

Estos dos argumentos hacen pensar que se está frente a la secuencia completa, y permitió confeccionar una secuencia consenso de 17.417 pb del gen de la miostatina bovina y sus regiones flanqueantes (Figura 32).

- Figura 32. Características finales de la secuencia ensamblada.



### 3 Estudio de la secuencia.

#### 3.1 Extracción ARN muscular esquelético.

La purificación de ARN muscular total se hizo según el protocolo de centrifugación en CsCl (Chirgwin y col., 1979) modificado. Debe usarse H<sub>2</sub>O DEPC para lavar todos los recipientes y hacer todas las soluciones utilizadas.

- Se hace una biopsia de tejido muscular y se congela rápidamente en nitrógeno líquido. Triturar el músculo sobre Nitrógeno líquido.
- Se transfiere el músculo triturado a un tubo falcon de 50ml. y se añade 5ml./500 mg. de solución de lisis.
- Mezclar vigorosamente con vortex.
- Romper el ADN pasando la mezcla varias veces por una aguja de jeringuilla.
- Poner en un tubo de 5 ml. de ultracentrifugación:
  - 1,5 ml. 5,7 CsCl (en H<sub>2</sub>O)
  - 3,5 ml. de la solución de lisis con el músculo.
- Ultracentrifugar a 35.000 r.p.m. (150.000g ) 16 h. a 18°C.
- Eliminar la solución de lisis y el CsCl.
- Cortar el tubo o limpiar sus paredes.
- Añadir 1 ml. de tampón de proteinasa K a temperatura ambiente.
- Pasar el contenido a un tubo falcon de 15 ml., e incubar a 50°C. para que se disuelva bien.
- Añadir 100 mg. de proteinasa K por ml. de solución.
- Incubar durante 2 h. a temperatura ambiente.
- Extraer 2 veces con fenol/cloroformo/alcohol isoamílico y una vez con cloroformo/isoamílico.
- Precipitar al Acetato de sodio pH=5,2 y etanol absoluto.
- Lavar con etanol 70%.
- Secar a temperatura ambiente.
- Resuspender en 50 µl. de H<sub>2</sub>O (DEPC)

#### 3.2 Estudio del ARN mensajero.

##### 3.2.1 La secuencia codificante de la miostatina está constituida por tres exones interrumpidos por dos intrones.

Para conocer la estructura del gen de la miostatina se alineó la secuencia genómica obtenida, con la secuencia codificante de la miostatina (Grobet y col., 1997). Este alineamiento confirmó la existencia de 2 intrones intercalados en la secuencia codificante. El intrón 1 se encuentra entre las bases 5.459-7.283, y el intrón 2 entre las bases 7659-9691. (ver Anexo3).

##### 3.2.2 Límites de los transcritos.

**3.2.2.1 Inicio de la transcripción.** El inicio de la transcripción se determinó a partir de ARN mensajero muscular, utilizando dos técnicas diferentes:

###### 3.2.2.1.1 C-RACE (Maruyama, 1995)



Es una técnica que sirve para identificar los posibles lugares de inicio de la transcripción, o lo que es lo mismo los posibles extremos del ARN mensajero en 5'.

El principio de esta técnica es el siguiente:

- 1- En la primera etapa se sintetiza un brazo de ADN codificante a partir de ARN total (GeneAmp RNA PCR Kit, Perkin-Elmer) utilizando el oligonucleótido C-RACE1

ARN	entre 1µg. y 5 µg.
Tampón PCR 10x	4 µl.
MgCl <sub>2</sub> (25mM.)	8 µl.
dATP (10mM.)	4 µl.
dGTP (10mM.)	4 µl.
dCTP (10mM.)	4 µl.
dTTP (10mM.)	4 µl.
Inhibidor de ARNasa	2 µl.
Transcriptasa inversa	2 µl.
C-RACE1 (5µM)	2 µl.
H <sub>2</sub> O DEPC	hasta 40 µl.

- Incubar 10' a temperatura ambiente
- 30' a 42°C.
- 5' a 99°C.
- 5' a 4°C.

Lisis del ARN, para dejar solamente ADN en la reacción

20 µl. de la reacción + 6,6 µl. NaOH 2M.  
incubar 10' a 37°C.

- Precipitar las dos reacciones:

26,6 µl. de reacción  
+ 2,8 µl. NaAc 3M., pH= 5,2  
+ 111 µl. etanol 100%  
guardar a -20°C. 30'  
centrifugar 15' a 10000 r.p.m. a 4°C.  
Lavar con etanol 70%.

- 2- El ADN de cadena simple se recirculariza y se utiliza como molde en una reacción de PCR con unos oligonucleótidos internos de la secuencia

Resuspender el producto anterior en 8 µl. de H<sub>2</sub>O

+ 1 µl. de tampón 10x  
+ 1 µl. de T<sub>4</sub> ARN ligasa

Incubar toda la noche a 22°C.

PCR utilizando como molde el ADN circularizado, utilizando los oligonucleótidos C-RACE2A y C-RACE-2B.

A continuación se realiza una PCR anidada (*Nested PCR*) con oligonucleótidos más internos, C-RACE-3A y C-RACE-3B.

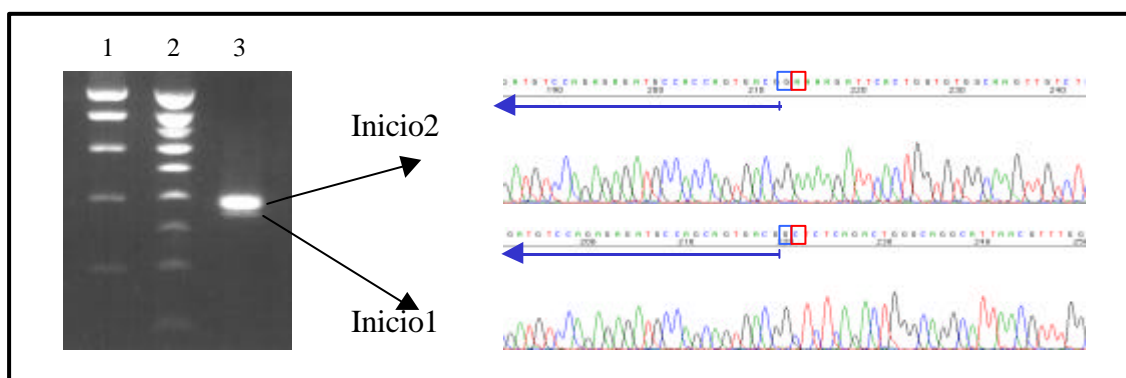
- Los productos resultantes se secuenciaron, y se alinearon las secuencias obtenidas sobre la secuencia conocida, para identificar el inicio de la transcripción (Figura 33).

- Tabla XIX. Oligonucleótidos utilizados en el protocolo de *C-RACE*.

Oligo	Secuencias
C-RACE1	5'-CGTCACTGCTGGCATCTCTCTGGACATCG-3'
C-RACE-2A	5'-GTCTTGAGGATGTAGTGTTTTCC-3'
C-RACE-2B	5'-GCCATAAAAATCCAAATCCTCAG-3'
C-RACE-3A	5'-CAAACATGCATTACACAGCCCCTC-3'
C-RACE-3B	5'-CTTCGCCTGGAAACAGCTCCTAAC-3'

Después de las dos rondas de PCR aparecen dos bandas (Figura 33), que indican la presencia de dos posibles lugares de inicio de la transcripción en el gen de la miostatina bovina en las posiciones 110 pb. (Inicio1) y 137 pb. (Inicio2) delante del codon de iniciación ATG (4.979 y 4.952 respectivamente, Anexo3; Figura 33). Subrayado se indica la secuencia que corresponde al oligonucleótido C-RACE-1. Por lo tanto la base siguiente, recuadrada en rojo, corresponderá con la posible base de inicio del transcrito de la miostatina bovina. En las secuencias de los dos posibles lugares de inicio de la transcripción aparece una G (recuadrado en azul, Figura 33) que no pertenece ni al oligonucleótido C-RACE1 ni a la secuencia de ADN genómico. Está descrita la posible incorporación de uno a varios nucleótidos provocado por la actividad terminal transferasa de la transcriptasa inversa (Superscript<sup>TM</sup>II) utilizada en este protocolo. La enzima puede añadir hasta 5 nucleótidos, fundamentalmente citosinas al encontrarse con la capucha que todo ARN completo tiene en posición 5' (Schmidt y Mueller, 1999).

- Figura 33. Gel de agarosa que muestra la presencia de dos posibles lugares de inicio del ARN mensajero de la miostatina bovina y su secuencia.



### 3.2.2.1.2 *Primer Extension* (Sambrook y col, 1989).

Es otro protocolo para identificar el inicio de la transcripción. Se basa en la amplificación de un oligonucleótido interno de la secuencia, cuya extensión empieza en algún lugar del ARN mensajero y acaba en el límite de este ARN mensajero.

Para ello se amplifica con un oligonucleótido marcado y se corre en un gel a la vez que una secuencia de ADN llevada a cabo con el mismo oligonucleótido, así se puede hacer corresponder el final de la extensión del oligonucleótido sobre el ARN mensajero con la base de la secuencia de ADN, e identificar a ésta como el inicio de la transcripción.

#### 1- Secuenciación (*T<sub>7</sub> Sequencing<sup>TM</sup> Kit*).

1.1 PCR sobre ADN genómico con dos oligos que flanqueen el inicio de la transcripción.- Se debe purificar el producto de PCR.

1.2 Acoplamiento del oligo.

- Desnaturalización: 64 µl. de producto PCR (30 ng./ µl.)  
16 µl. NaOH 2M.  
Vortex y centrifugar  
Incubar 10' a temperatura ambiente  
+ 14 µl. NaAc 3M.,  
+ 8 µl. H<sub>2</sub>O  
+ 240 µl. etanol 100% mezclar bien y mantener 15' en nieve carbónica.  
Centrifugar a 4°C. 20' a 13.000 r.p.m.  
Eliminar el sobrenadante  
Lavar con etanol 70%.  
Centrifugar 10' y eliminar el sobrenadante.  
Dejar secar, sin utilizar vacío.  
Resuspender en 20 µl. de H<sub>2</sub>O.
- Hibridación del oligo.

Añadir a los 20 µl.:  
+ 4 µl. del oligo (5 pmol./ µl.)  
+ 4 µl. de tampón de *annealing*.  
Vortex ligero y centrifugar.  
Incubar 5' a 65°C.  
Incubar 10' a 37°C.  
Mantener los tubos a T<sup>a</sup> ambiente ≥ 5'.

1.3 Secuenciación (*Labeling mix* dATP).

Diluir el enzima T<sub>7</sub> ADN polimerasa.: 6 µl. T<sub>7</sub> ADN polimerasa  
24 µl. tampón de dilución frío.

Preparar 4 tubos por reacción:

5 µl. "A" <i>mix short</i>
5 µl. "C" <i>mix short</i>
5 µl. "G" <i>mix short</i>
5 µl. "T" <i>mix short</i>

Reacción de marcaje (por cada dNTP):	<i>Labeling mix A</i>	6 µl.
	$\alpha^{35}$ dATP	2 µl.
	T <sub>7</sub> ADN polimerasa diluida	4 µl.

Mezclar suavemente e incubar 5' a temperatura ambiente.  
Durante estos 5' calentar los tubos de los dNTP a 37°C. al menos 1'.

1.4 Terminación. (debe ser inmediatamente después de los 5').

Transferir 9 µl. de la reacción de marcaje a cada tubo con dNTP.  
Mezclar con la pipeta.  
Incubar 5' a 37°C.  
+ 5 µl. solución de STOP  
Guardar a 4°C.

## 2- Extensión del oligo.

### 2.1 Marcado del oligo.

oligo (33 µM.)	1 µl.
Tampón <i>forward</i> 5x	14 µl.
$\chi^{32}$ ATP	15 µl.
T <sub>4</sub> kinasa	8 µl.
H <sub>2</sub> O DEPC	32 µl.

Incubar a 37°C. 1 hora.  
Parar la reacción: 2' a 90°C.

### 2.2 Hibridación del oligo marcado.

Oligo marcado	1 µl.
ARN	10 µg.
NaAc 3M. DEPC	2 µl.
Etanol 100% DEPC	45 µl.

Mantener a -20°C. durante 30'.  
Centrifugar 20' en frío.  
Lavar con etanol 70% DEPC.  
Centrifugar 20' en frío.  
Eliminar el sobrenadante.  
Dejar secar a temperatura ambiente, no usar vacío.  
Resuspender en 30 µl. de tampón de hibridación.  
Mezclar vigorosamente con la pipeta.  
Incubar a 60°C. durante 10'.

Incubar a 80°C. durante 10'.  
Hibridar a 50°C. durante 8-16 h.

Precipitar el producto de hibridación: +170 µl. H<sub>2</sub>O DEPC  
+400 µl. etanol 100%

Mezclar y mantener en hielo 1h.

Centrifugar 30'.

Lavar con etanol 70% DEPC.

Eliminar el sobrenadante y evaporar la muestra a temperatura ambiente.

Resuspender en 20 µl. de 1x <i>R.T. Buffer</i> :	5x <i>first strand buffer</i>	4 µl.
	DTT 0,1M.	2µl.
	dNTP 10mM.	1µl.
	H <sub>2</sub> O DEPC	13µl.

+ 1 µl. de Superscript II (SUPERScript™ II, Gibco-BRL) y mezclar.

Incubar 1 h. a 42°C.

Inactivar la enzima a 70°C. durante 15'.

+ 1 µl. de ARNasa, e incubar 20' a 37°C.

+ 150 µl. T<sub>10</sub>E<sub>1</sub> pH=7,6.

+200 µl. fenol:cloroformo:alcohol isoamílico (25:24:1)

Vortex durante 30'' y centrifugar 5'.

Recoger el sobrenadante en otro tubo, añadir 500 µl. de Etanol 100% y mantener 1h. en hielo.

Centrifugar 15' y eliminar el sobrenadante.

Lavar con etanol 70%

Centrifugar 15', eliminar el sobrenadante y dejar secar a temperatura ambiente.

Guardar a -20°C.

## 2- Migración de los productos en gel de acrilamida.

En un gel de acrilamida 6% (acrilamida/bisacrilamida 19:1) (Gibco-BRL), se cargan las 4 reacciones de la secuenciación, y en pocillo contiguo la reacción de extensión.

Los resultados obtenidos por esta técnica no fueron concluyentes, y no se tuvieron en cuenta.

### 3.2.2.2 Final de la transcripción.

Los finales de la transcripción, se localizaron utilizando la técnica de 3'RACE. Brevemente en este protocolo se sintetiza un brazo de ADN codificante de la miostatina a partir de ARN mensajero utilizando un oligonucleótido poli-T y un oligonucleótido interno (A) de la miostatina bovina. Posteriormente se lleva a cabo otra PCR ahora anidada cambiando el oligonucleótido interno por otro más cercano al posible final de la transcripción (B). Como no se conoce bien donde puede estar este final se utilizaron 3 protocolos diferentes para intentar no dejar ningún posible lugar fuera del análisis. Se utilizaron los oligonucleótidos de la Tabla XX.

1 Síntesis del ADN codificante. En un tubo de 1,5 ml. libre de ARNasas,

ARN total	3 µg.
Oligo poli-T + adaptador (10 µM.)	1 µl.
H <sub>2</sub> O DEPC	hasta 12 µl.

- Incubar a 70°C. durante 10', meter en hielo >1' y centrifugar.

Tampón de PCR 10x	2 µl.
MgCl <sub>2</sub> (25 mM.)	2 µl.
dNTP (10mM.)	1 µl.
DTT 0,1 M.	2 µl.

- Mezclar y centrifugar

- Incubar a 42°C. durante 5'.

SuperScript II RT (retrotranscriptasa)	1 µl.
--	-------

- Incubar a 42°C. durante 50'.

- Terminar la reacción incubando a 70°C. durante 15'.

- Poner en hielo y centrifugar.

- El primer brazo de ADN codificante ya ha sido sintetizado con la retrotranscriptasa.

ARNasa H	1 µl.
----------	-------

- Incubar a 37°C. durante 20'.

- Con esto se digiere el ARN y solo queda el ADN codificante sintetizado, que se puede guardar a -20°.

1- Amplificación del ADN codificante.

Tampón PCR 10x	5µl.
MgCl <sub>2</sub> (25mM.)	3 µl.
dNTP (10mM.)	1 µl.
oligo interno (10 µM.)	1 µl.
Oligo adaptador (10 µM.)	1 µl.
Taq polimerasa	0,5 u.
ADN codificante.	2 µl.
H <sub>2</sub> O	hasta 50 µl.

Programa de PCR (P.E. 9600)

1x	94°C. 5'
35x	94°C 30''
	55°C. 1'
	72°C. 1'.
1x	72°C. 10'.

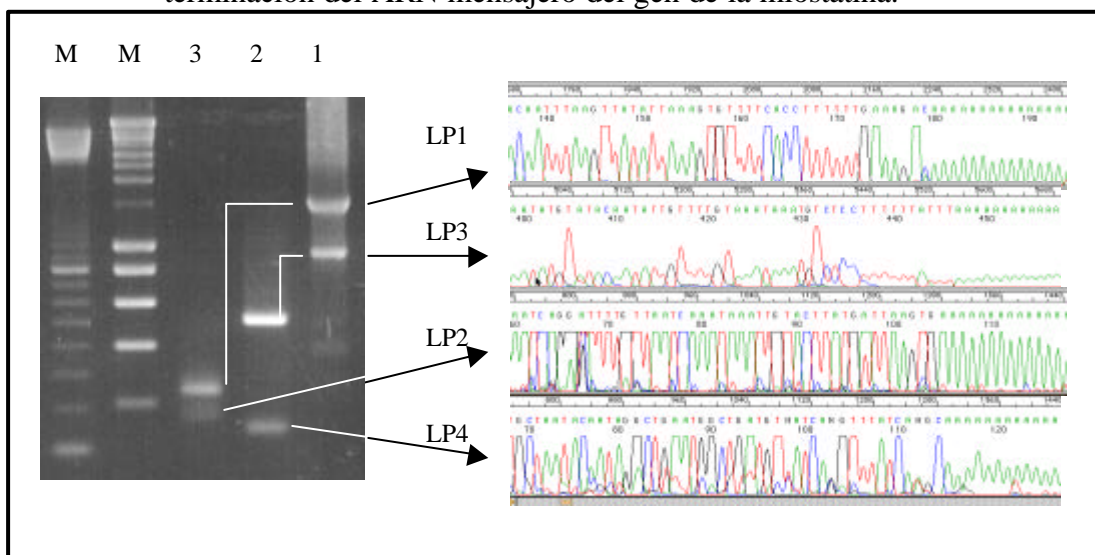
- 2- *Nested* PCR: siguiendo el mismo protocolo, pero cambiando el oligo por uno más interno, y en vez de ADN codificante se añaden 2  $\mu$ l. del producto de PCR anterior.
- 3- Clonado de los productos de RACE3': Los productos de PCR se visualizan en gel de agarosa, y se recuperan las bandas obtenidas. Se extrae el ADN de cada una de ellas (protocolo *GeneClean*), y se clonan en T-A (*Topo TA Cloning<sup>®</sup> Kit*, Invitrogen<sup>®</sup>).
- 5- Secuenciación de los productos de *RACE3'* clonados, según el protocolo ABI PRISM<sup>™</sup> *Dye Terminator Cycle Sequencing, Ready reaction kit*, en un secuenciador automático ABI 373, Perkin Elmer.

- Tabla XX. Oligonucleótidos utilizados en el protocolo de 3'-*RACE*.

Oligo	Secuencias
5-RACE-1A	5'-CCCCAGAGGTTTCAGCCGGCCCTGC-3'
5-RACE-1B	5'-CCAGCCATGGTAGTAGATCGCTGTGGG-3'
5-RACE-2A	5'-ATTCCACAAAGTAGGGATGGCACAC-3'
5-RACE-2B	5'-TACGCAGGTTACCATTCTTACTG-3'
5-RACE-3A	5'-TTTGGTATATTTTACAGTAAGGAC-3'
5-RACE-3B	5'-TAAACTGATGATATCTTACAAATTG-3'

En la Figura 34 se puede ver un gel de agarosa con los posibles lugares de finalización de los transcritos de la miostatina bovina obtenidos con los tres protocolos de PCR descritos. A la derecha están situadas las secuencias resultantes. Algunos de los lugares de terminación del ARN mensajeros se detectaron en dos protocolos diferentes (LP1 y LP3). El final de la transcripción corresponde con la base situada inmediatamente anterior a la cola poliA en las secuencias de ARN mensajero. Los lugares de poliadenilación se nombran en el orden en que se encuentran sobre la secuencia. En algunos lugares la presencia de Adeninas en la secuencia hacen imposible identificar la base exacta de terminación del transcrito, por eso se indican varias bases como posibles finales de la transcripción. Los posibles lugares de finalización del transcrito de la miostatina están situados en las posiciones: LP1 11578-82; LP2 11507-10; LP3 11028-9; LP4 10681-7 (Anexo3).

- Figura 34. Gel de agarosa donde se muestra la presencia de los 4 posibles lugares de terminación del ARN mensajero del gen de la miostatina.



## Capítulo 3.

### Identificación de las posibles regiones reguladoras del gen de la miostatina.

---

#### 1 Antecedentes.

Los mecanismos que regulan la expresión del gen de la miostatina se desconocen. Comparando las secuencias de la miostatina del ratón (L. Grobet y V. Marot, comunicación personal) y bovina, que comprenden el gen completo y parte de las regiones flanqueantes, y basándose en la conservación de secuencias, van a intentar identificarse las regiones que pueden participar en el control de la transcripción del gen de la miostatina.

#### 1.1 Identificación de regiones reguladoras de la transcripción.

El término de huellas filogenéticas (*phylogenetic footprinting*) describe las comparaciones filogenéticas que identifican elementos conservados en *cis*, en zonas no codificantes de genes homólogos. En sentido más amplio se denomina así a la identificación de cualquier región funcional utilizando la comparación de secuencias de segmentos genómicos ortólogos (Fickett y Wasserman, 2000). Una cuestión importante es conocer cual es la distancia óptima en términos de evolución entre las especies que se comparan. Generalmente la mayoría de las regiones reguladoras humanas pueden ser identificadas si se comparan las secuencias humanas y de roedores, pero puede variar dependiendo del gen en cuestión (Fickett y Wasserman, 2000). Cuando el patrón de expresión del gen es diferente entre las dos especies estudiadas, suele ser debido a cambios en los elementos reguladores. Entonces en estos casos las especies elegidas deben estar más relacionadas para que las comparaciones sean informativas. En el caso de otros *loci* la tasa de mutación en esa región puede ser tan baja que las comparaciones humano-ratón no sean informativas. En ese caso se necesitarán alineamientos de secuencias de especies más distantes (Hardison y col., 1997).

Pero las huellas filogenéticas no dan una información de la función de las regiones conservadas. Para predecir la funcionalidad de esas regiones es necesario un análisis de los factores de transcripción que pueden unirse específicamente a ellas. Normalmente los factores de transcripción se identifican por medio de una secuencia consenso, sin embargo una matriz que pondere la importancia de cada base en cada posición del módulo (*position weight matrix*) es más apropiado. Esta matriz asigna a cada nucleótido en cada posición un peso específico, reflejando la frecuencia de ese nucleótido en esa posición. El valor de un determinado elemento se obtiene de la suma de los correspondientes pesos. De esta forma se obtiene más información que a partir de una secuencia consenso (Fickett y Wasserman, 2000).



La expresión del gen de la miostatina en ratón se detecta ya en el día 9,5 dentro de los somitos, más específicamente en el miotomo. A medida que avanza el desarrollo embrionario, el ARN mensajero de la miostatina se detecta en un gran número de músculos (McPherron y col, 1997). Los niveles de ARN mensajero, según un trabajo realizado en la especie porcina, aumentan y alcanzan un máximo durante las últimas fases de la gestación y disminuyen después del nacimiento, sin embargo se sigue encontrando ARN mensajero de miostatina en tejido muscular esquelético adulto. También se ha descrito la presencia de ARN mensajero en tejido muscular cardíaco y en glándula mamaria (Ji y col., 1998). En músculo esquelético en fase de regeneración, la miostatina se detecta en las células mononucleadas situadas en la periferia de las fibras musculares.

La mayoría de las regiones reguladoras, incluyendo los promotores, están influenciados en algún modo por el contexto específico espacio-temporal donde se encuentran. El análisis de regiones reguladoras conocidas, indica que muchos elementos músculo-específicos están concentrados en módulos reguladores de 200 pb. Para identificar regiones con probabilidad de conferir una expresión músculo-específica a los genes, se pueden utilizar las huellas filogenéticas complementado con algoritmos de regresión logística. Dentro de las regiones conservadas, encontradas por huellas filogenéticas, se buscarán elementos en *cis* que se sabe que confieren especificidad de expresión en tejido muscular. Este análisis se llevará a cabo utilizando un algoritmo de regresión logística que se basa en la asunción de que las regiones reguladoras de la expresión específica muscular están compuestas de grupos de diferentes factores de transcripción musculares incluidos dentro de regiones de 200 pb. de tamaño. En este análisis se consideraron los siguientes factores de transcripción: Tef, Sp-1, MEF-2, Myf y SRF (Wasserman y Fickett, 1998), para los que existen ya desarrolladas matrices de pesos por posición de cada uno de los nucleótidos (Wasserman y Fickett, 1998; <http://agave.humgen.upenn.edu/MTIR/HomePAge.html>).

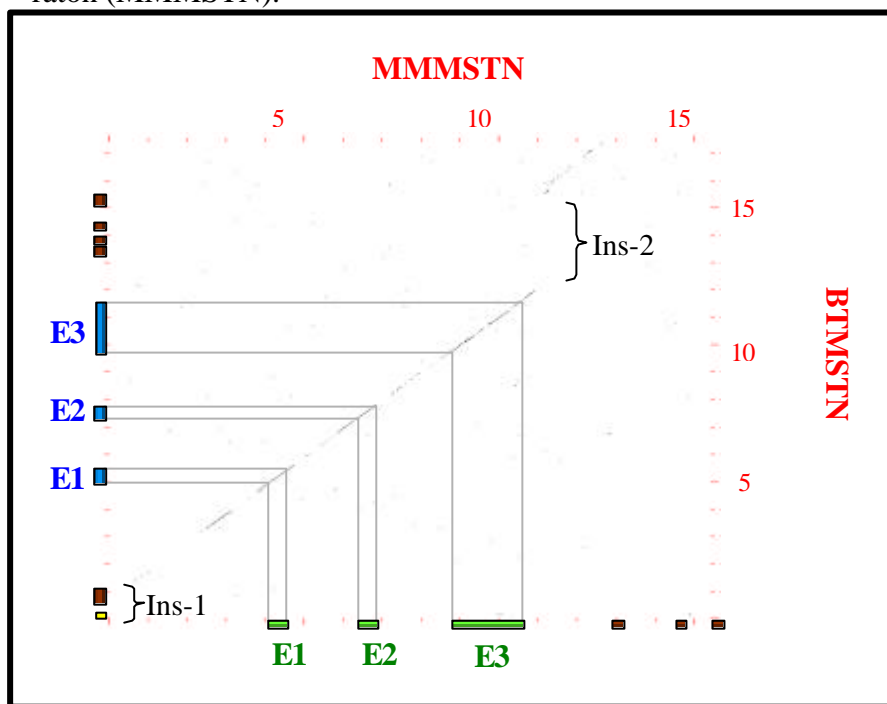
## 2 Comparación de las secuencias de la miostatina bovina y de ratón.

Se compararon las secuencias de la miostatina bovina y de la de ratón (L. Grobet y V. Marot) utilizando una matriz de puntos realizada con el programa *Compare* y representada gráficamente con el programa *DotPlot* del grupo de programas *GCG* (Figura 35).

El programa *Compare*, compara dos secuencias de ácidos nucleicos o proteínas y crea un archivo de puntos de similitud entre ellas, que se puede representar gráficamente con el programa *DotPlot*. Las secuencias se comparan en cada una de todas las posibles localizaciones y crea un punto donde las secuencias son similares, las coordenadas de esos puntos se almacenan en el archivo de salida. El criterio que se utiliza para indicar que dos secuencias son similares es decisión de cada uno. Normalmente el programa compara dos secuencias en cada registro, buscando en todos los lugares el número de identidades que aparecen en una ventana de un tamaño dado, para crear un punto cuando se alcanza el número suficiente de nucleótidos iguales. Este criterio se puede decidir de antemano, el tamaño de la ventana y el número de nucleótidos iguales necesarios para situar un punto.

El programa *DotPlot*, representa la homología de dos secuencias en un mapa de puntos utilizando la salida del programa *Compare*. Este método es el más útil para ver las estructuras comunes entre dos secuencias, o las estructuras repetitivas dentro de una sola secuencia. Este programa es la segunda parte de un grupo de programas que generan imágenes de puntos de similitud entre dos secuencias (Maizel y Lenk, 1981). *Compare* crea un fichero de salida con las coordenadas de cada punto en común entre dos secuencias y *DotPlot* plasma estos puntos en un gráfico.

- Figura 35. Comparación de las secuencias bovina (BTMSTN) y de ratón (MMMSTN).



La similitud comienza aproximadamente a partir de la base 1.300 de la secuencia bovina y 500 de la secuencia de ratón. A partir de entonces se observa una línea diagonal, indicando la homología de las dos secuencias. Esta línea se corta aproximadamente en la posición 12.330 pb. de la secuencia bovina y 10.800 pb. de la secuencia de ratón, volviendo a aparecer en las posiciones 15.850 bovina y 11.200 del ratón. La secuencia de ratón se extiende 2.700 pb. más, no encontrándose obviamente ninguna homología. En la secuencia bovina se puede diferenciar dos zonas sin homología con la de ratón, que se identifican como Inserción 1 e Inserción 2 (Figura 35-36), aunque pueda tratarse del proceso contrario es decir delección de los fragmentos identificados en la secuencia de ratón. Sería necesario el análisis de la secuencia de una tercera especie para poder aclarar este proceso.

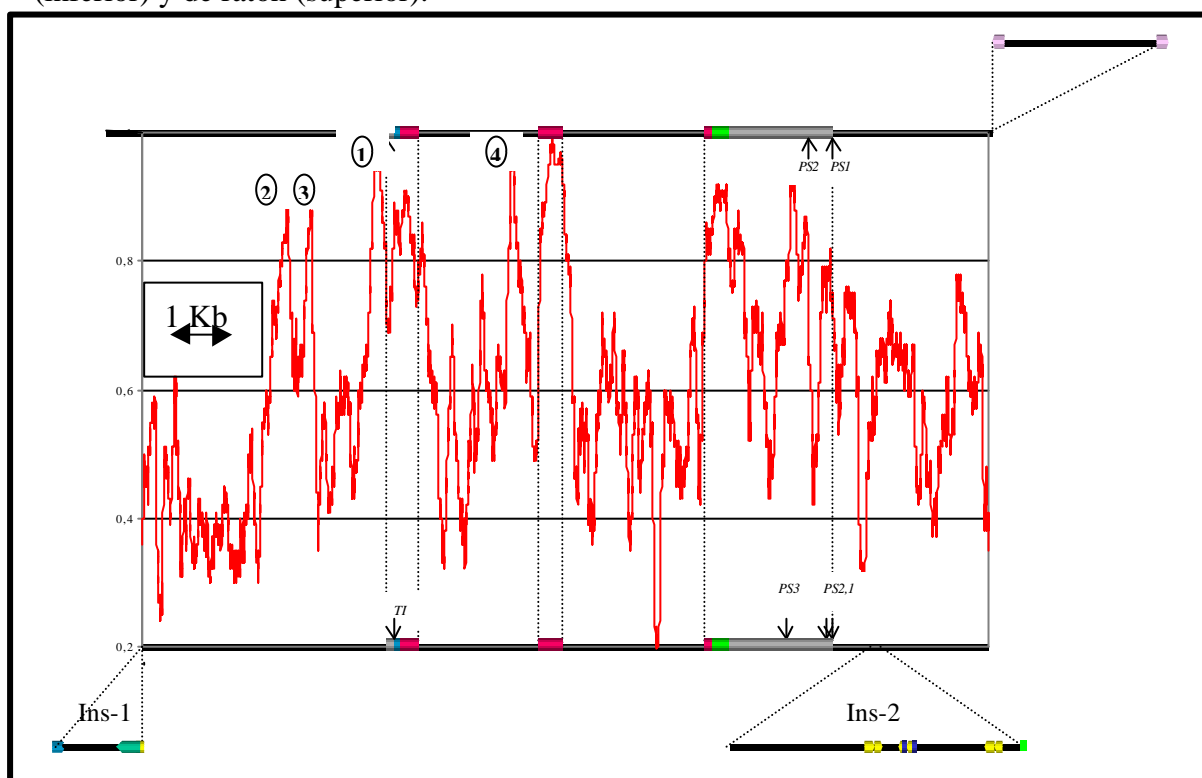
En un posterior análisis se eliminaron las secuencias que quedaban fuera de la región que mostraba homología, identificadas en la Figura 35-36 como Ins-1 e Ins-2. Las secuencias, bovina y de ratón, resultantes se alinearon usando el programa *Gap* del grupo de programas *GCG* (ver Figura 36).

El programa *Gap* encuentra similitudes entre dos secuencias completas usando el algoritmo de Needleman y Wunsch (Needleman y Wunsch, 1970). Considera todos los

posibles alineamientos y huecos entre las dos secuencias y crea un alineamiento definitivo que maximiza el número de coincidencias y minimiza el tamaño y número de los huecos.

El resultado fue un alineamiento de 12.772 caracteres, que incluía 12.477 pb. de la secuencia bovina y 12.039 de la de ratón. La secuencia de ratón contenía 295 pb. que no tenía la bovina (12.772-12.477), mientras que la bovina contenía un segmento de 438 pb. que no aparecía en la de ratón (12.477-12.039). El alineamiento obtenido se analizó con una ventana de 100 pb. calculando la similitud como el porcentaje de homología entre las dos secuencias en cada ventana.

- Figura 36. Alineamiento de las secuencias homólogas de la miostatina bovina (inferior) y de ratón (superior).



Se observó un alto porcentaje de homología en las regiones de los exones, siendo mayor en la segunda mitad del exón 2.

En la zona en 5', justo antes del inicio de la transcripción, se encuentra otro segmento altamente conservado. Este segmento contiene una TATA-box y una CAAT-box, con lo cual podemos afirmar, con casi total seguridad, que debe tratarse de la región del promotor (Zona Conservada 1; ver Discusión 2.2, Figura 39).

Se encuentran otras regiones en 5' donde hay también una alta homología de secuencias, identificadas como Zona Conservada 2 y 3. En el tercio posterior del intrón 1 se localiza la última región de secuencia conservada, que se identifica como Zona Conservada 4 (Figura 36).

En la zona en 3' no traducida también se observaron unas zonas de alta conservación de secuencia:

- Una primera zona justo después del codón stop.
- Una segunda zona que correspondería con los alrededores del tercer posible lugar de poliadenilación (LP) bovino y segundo de ratón.
- La tercera correspondería a la coincidencia entre el LP2 y LP1 bovinos y el LP1 murino.

### **3 Búsqueda de regiones reguladoras músculo-específicas.**

En esta búsqueda se consideraron los siguientes factores de transcripción, conocidos por conferir especificidad de expresión muscular: Tef, Sp-1, MEF-2, Myf y SRF, utilizando algoritmos de regresión logística (Wasserman y Fickett, 1998).

Con los métodos utilizados para la búsqueda de factores de transcripción músculo-específicos, no se logró encontrar ninguno de ellos dentro de las zonas evidenciadas mediante las huellas filogenéticas, es decir las Zonas Conservadas 2, 3 y 4. Sin embargo se piensa que esas zonas evidenciadas, con alta homología de secuencias, localizadas en 5' del punto de iniciación de la transcripción y en el intrón 1 (Figura 35 y 36), deben tener alguna función dentro de la regulación del gen.

## Discusión.

---

### 1- Reducción del intervalo cromosómico:

El mayor esfuerzo realizado en los últimos años en genética molecular aplicada a la producción animal ha sido para localizar o mapear *loci* de caracteres con importancia económica en producción animal utilizando el análisis de ligamiento (Georges 1998). Estos datos pueden llegar a utilizarse directamente mediante estrategias de MAS (*Marker Assisted Selection*; Visscher y col., 1998), sin embargo la culminación de estos trabajos requiere la identificación y el clonado posicional de los genes causantes (Anderson 1998).

Para intentar el clonado posicional de genes de interés económico en producción animal, se puede aprovechar el conocimiento mayor de los genomas humano y de ratón. Se ha estimado que el tamaño de los segmentos de cromosomas conservados entre especies es de unos 10 cM de media entre humano y ratón (Copeland y col., 1993) y posiblemente mayor entre humano y las especies ganaderas. La identificación de fragmentos conservados y sus límites en las regiones de interés va a facilitar las estrategias de clonado de candidatos posicionales utilizando los mapas de transcritos humanos y de ratón (Archibald 1998).

Lo habitual en un trabajo de clonado posicional de un gen de interés económico es intentar reducir lo más posible el intervalo cromosómico donde se localiza el *locus* y lo primero es saturar la región candidata con marcadores, hasta llegar a un intervalo cromosómico lo más pequeño posible donde poder iniciar la construcción de un mapa físico. Como se ha visto en el capítulo 1, con la intención de reducir la región que contiene al *locus mh* (2q11-21), se aislaron 5 marcadores microsatélites nuevos en esta región.

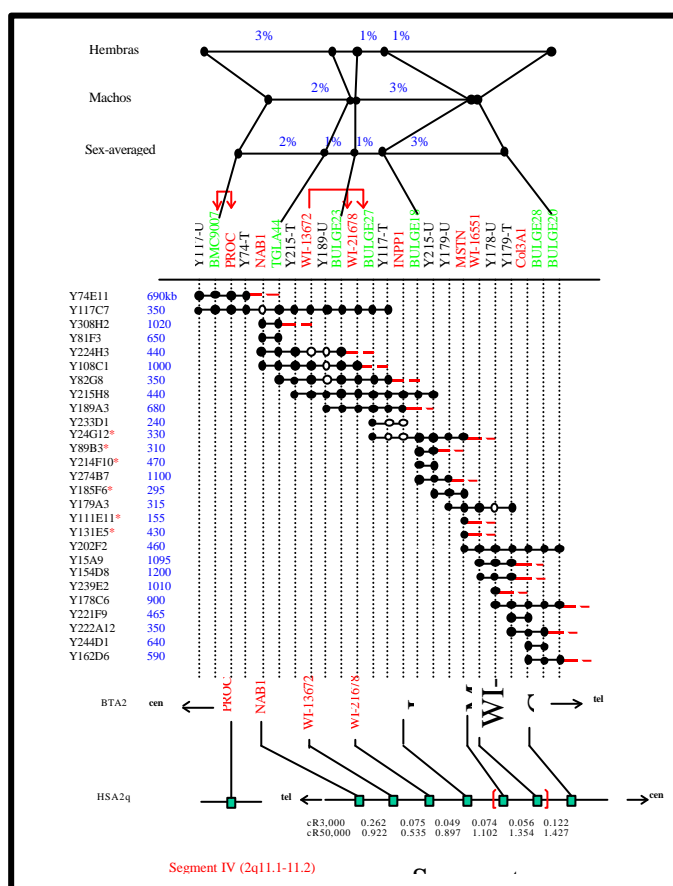
A partir del conocimiento de la región bovina donde se localizaba el gen (2q11-21) de la miostatina, se construyó un *contig* de YAC ayudándose de las etiquetas (STS) mapeadas en esa región. Se utilizaron 7 marcadores tipo microsatélites (BMC9007, TGLA44, y los 5 resultado de este trabajo: Bulge23, Bulge27, Bulge18, Bulge28, Bulge20), 8 genes (EST), y 9 STS correspondientes a extremos de YAC. El *contig* se componía de 27 YAC con una profundidad media de 4,3 YAC/etiqueta, y se estimó que comprendía 1,2 Mb. de la región centromérica del cromosoma 2 bovino (Pirottin y col., 1999).

Con los 5 microsatélites que resultaron polimórficos, se construyó un mapa de ligamiento, que mide de extremo a extremo del *contig* 7 cM. (BMC9007-BULGE20), lo que supone una relación de 160 kb./cM. en esta región estudiada (Pirottin y col., 1999).

El mapeo comparativo permitió mostrar que el *contig* incluía una zona de rotura cromosómica que unía dos segmentos de cromosoma sinténicos pero no adyacentes en humano. Comparando el orden de los genes ortólogos humanos de la región, verificado mediante el uso de un panel de híbridos de irradiación del genoma humano completo, y el orden bovino inferido a partir del *contig*, se confirmó la conservación del orden de los genes dentro de los segmentos cromosómicos.

La aparición de un gen candidato posicional del *locus mh*, que finalmente permitió la identificación de la miostatina bovina como gen responsable de la hipertrofia muscular, hizo que la estrategia clásica de un clonado posicional se viera trastocada. Los microsatélites nuevos, utilizados en análisis de ligamiento, permitieron acortar el intervalo donde buscar genes candidatos (ver Capítulo 1 y Pirottin y col., 1999).

- Figura 37. *Contig* de YAC de la región bovina de la miostatina (Pirottin y col., 1999).



## 2- Estructura del gen de la miostatina bovina.

### 2.1 Transcritos de ARN mensajero de la miostatina bovina.

La secuencia codificante de la miostatina está constituida por tres exones interrumpidos por dos intrones, evidenciado al alinear la secuencia completa (Capítulo 2) con la secuencia codificante (Grobet y col., 1997).

a) Intrón 1. El intrón 1 se localiza entre las bases 5460 y 7285 (Anexo3) y cumplen la regla del GT-AG. Se puede decir que el intrón 1 es un intrón de fase 1, ya que interrumpe un codon serina entre la base uno y dos. En la región en 3' del intrón se encuentra una zona rica en pirimidinas y a -67 pb. del exón 2. En ella hay una adenina situada en un consenso apropiado (UGAU, 7.216-20, Anexo3) para constituir el punto de formación del lazo, en el proceso de corte y empalme de los intrones (Keller y Noon, 1984).

b) Intrón 2. Los límites identificados siguen la regla GT-AG (7660-9.691, Anexo3). Este intrón es de fase 0, ya que interrumpe la secuencia codificante entre dos codones adyacentes. La zona inmediatamente anterior al extremo 3' del intrón se caracteriza por una secuencia rica en pirimidinas. Existe una adenina en un contexto apropiado (UUAAC, 9.666-9.670, Anexo3) situada a -23 pb. del límite 3' del exón, constituyendo el posible lugar de formación del lazo para la maduración del ARN mensajero por corte y empalme (Keller y Noon, 1984).

Se han identificado dos posibles lugares de inicio de la transcripción, situados aproximadamente 110 pb. (Inicio1) y 137 pb. (Inicio2) delante del codón de iniciación (4.979 y 4.952 respectivamente, Anexo3). No se encontró ninguna secuencia consenso de iniciación (YYANWYY) en estos lugares de iniciación. El sitio Inicio1 corresponde aproximadamente al lugar de iniciación identificado en ratón por McPherron y col., 1997.

La secuenciación de los productos de PCR obtenidos por la técnica de *RACE3'* identificó 4 posibles lugares de poliadenilación (LP: Lugar de Poliadenilación) situados a 612 pb. (LP4), 959 pb. (LP3), 1.438 pb. (LP2) y 1.509 pb. (LP1) a partir del codón stop (10.681-7, 11.028-9, 11.507-10 y 11.578-11.582 respectivamente, Anexo3). El tamaño de los productos PCR encaja con la secuencia del ADN genómico, lo que indica que no existen fenómenos de procesamiento adicionales en el extremo 3' no traducido. Resultó imposible determinar el lugar exacto de terminación del ARN mensajero, debido a la presencia de adeninas en la secuencia. Por ejemplo, en los fragmentos correspondientes a LP1, la mitad de ellos contienen una citosina en el límite entre el ADN y la cola poli-(A), 1.511 pb. después del codón stop. Lo más probable es que sea debido a un mal acoplamiento del oligo poli-(T) usado en el protocolo para la síntesis del ADN codificante, a causa de las adeninas presentes en la secuencia.

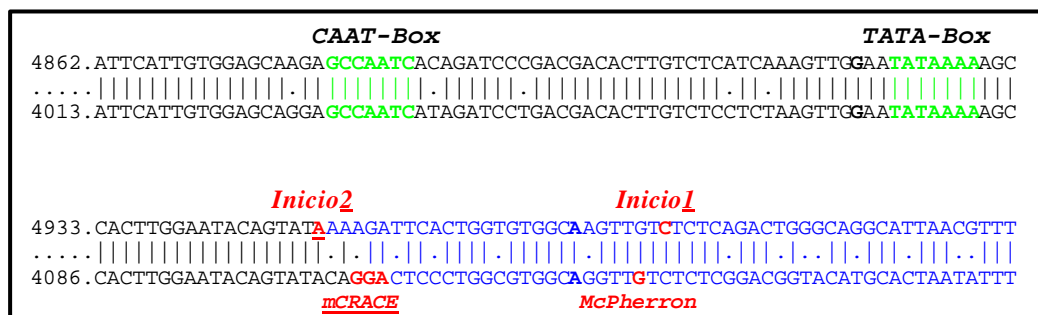
La señal de poliadenilación AAUAAA precede en 19 y 17 pb. respectivamente a los dos lugares de poliadenilación LP2 y LP3, y la señal AUUAAA al LP1 en 23 pb. Sin embargo no se encuentra una señal de este tipo delante del lugar de poliadenilación LP4. Además este lugar esta precedido de una serie de 6 adeninas, por lo que puede deberse a un





Para caracterizar el promotor del gen de la miostatina se alinearon las secuencias bovina y de ratón, localizadas inmediatamente en 5' del inicio de transcripción identificado (Figura 39).

- Figura 39. Alineamiento en la región del promotor, y localización de la TATA-box y CAAT-box.



- En azul la secuencia del promotor mínimo de la miostatina bovina, y en rojo, el promotor de la miostatina de ratón. Las cajas TATA-box y CAAT box (en verde) se localizan en regiones con una gran homología de secuencias..

El promotor de la miostatina es un promotor dirigido por una TATA-box, situada a -21 pb del punto de inicio2 de la transcripción (4.925-4.931, Anexo3). Además contiene una CAAT-box a -65 pb. del punto de inicio2 (4.881-4.887, Anexo3), situadas en una región con una alta homología de secuencias. En una región de unos 80 pb se encontraron los dos posibles lugares de inicio de la transcripción, una TATA-box y una CAAT-box, con lo que todo hace indicar que se está frente al promotor de la miostatina. Además esta región ha sido identificada como un segmento de secuencia altamente conservado (Zona Conservada1; Figura 36).

### 3 Estudio de la proteína.

El codon de iniciación de las proteínas eucariotas es AUG. El codon de iniciación de la miostatina fue el primero que se encontró en la secuencia después del inicio de la transcripción. El contexto de este codon AUG no coincide con el contexto óptimo de Kozak (GCCACCatgG), pero sí con un contexto adecuado (RNNatgY), donde R es una purina e Y es una pirimidina (Kozak, 1996). En el caso de la miostatina bovina se encontró una secuencia ACCatgC (nucleótidos 5084-5090, Anexo3). Además utilizando este inicio se obtenía un marco de lectura abierta que traducido, formaba una proteína con las características básicas de la familia de factores de crecimiento TGF- $\beta$ . Todo esto hace pensar que el codon identificado es el verdadero codon de inicio de la traducción.

Basándose en el análisis de la secuencia aminoacídica formada a partir del codon AUG predeterminado, se constató que la miostatina codificaba para una proteína con las características típicas de un miembro de la familia TGF- $\beta$ . Estaba formada por un péptido señal, un péptido latente, y un dominio carboxi-terminal caracterizado por la existencia de nueve residuos cisteína, teóricamente involucrados en la formación de puentes disulfuro

intra- e intercatenarios. Teniendo en cuenta el resto de miembros de la familia TGF- $\beta$  y por analogía con éstos, se supuso que la miostatina se sintetizaba como una preproteína, que era procesada proteolíticamente, para dejar libres las moléculas bioactivas, que formarán dímeros del dominio carboxi-terminal, unidos por puentes disulfuro (Massague, 1992; McPherron y Lee, 1996).

Tras analizar la secuencia de la proteína con el método de von Heijne (1987), se encontró un residuo de 23 aminoácidos, que formaban el péptido señal, inmediatamente después del primer residuo metionina (MQKLQISVYIYLFMLIVAGPVDL, aminoácidos 1-23). En el péptido latente se encontraron 2 posibles lugares de N-glicosilación (NTT, 47-49; NIS, 71-73), 4 residuos cisteína en las posiciones 39-42-137-138, y una diana de proteólisis (RSRR, aminoácidos 263 al 266) que liberaba el fragmento carboxi-terminal bioactivo. Dentro de este fragmento carboxiterminal se encontraron nueve residuos de cisteína (272-281-282-309-313-339-340-372-374) y uno de glicina (311) localizados en las posiciones típicas conservadas de los miembros de la familia TGF- $\beta$ . Además en el péptido latente los cuatro residuos de cisteína localizados, podrían estar relacionados con puentes disulfuro intra- o intermoleculares (ver Anexo3).

### 3.1 Estructura molecular de la miostatina.

La miostatina es una proteína que pertenece a la familia TGF- $\beta$ , factores de crecimiento que controlan la proliferación y diferenciación celular, modulando la síntesis de matriz extracelular y la función biológica de otros factores de crecimiento. Parece que los TGF- $\beta$  juegan un papel muy importante en la homeostasis de los tejidos y en la respuesta inmunitaria.

Los miembros de la familia TGF- $\beta$  se sintetizan como largas moléculas precursoras, pre-pro-proteínas; con un péptido corto señal, una zona aminoterminal y la zona carboxiterminal bioactiva. Su estructura típica (McPherron y Lee, 1996) se caracteriza por:

- Un grupo de aminoácidos hidrofóbicos cerca del extremo N-terminal que funciona como señal para la secreción.
- Una forma precursora que es proteolíticamente procesada, liberando un fragmento carboxiterminal bioactivo y otro aminoterminal. Al fragmento aminoterminal se le denomina péptido asociado latente ( $\beta$ -LAP), y mientras permanece unido, de forma no covalente, mantiene al TGF- $\beta$  en un estado latente.
- Una región carboxiterminal muy conservada entre los miembros de la familia, que contiene 9 residuos de cisteína localizados en lugares característicos. La secuencia aminoácídica del péptido bioactivo presenta un 70-80 % de homología entre los diferentes miembros de la familia, llegando hasta un 90 % de identidad entre la miostatina y GDF-11 (McPherron y col., 1999). En las

diferentes especies la secuencia de aminoácidos de cada uno, es prácticamente idéntica.

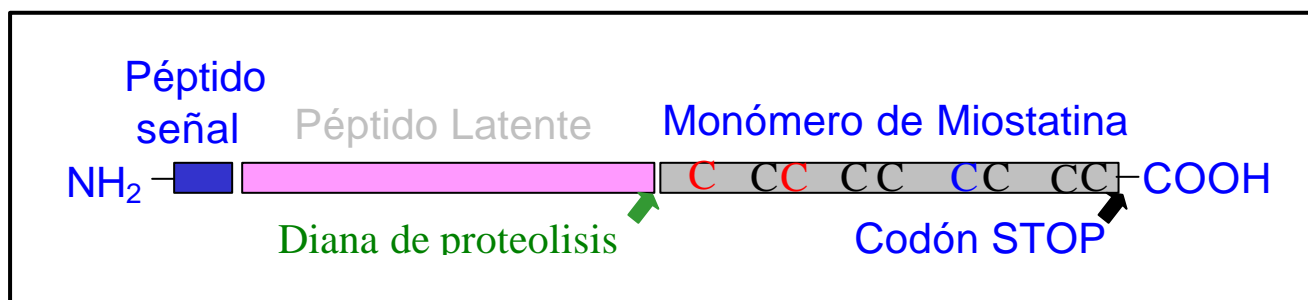
- La molécula bioactiva está formada por un dímero de fragmentos carboxiterminales unidos por puentes disulfuro. En la mayoría de los casos esta forma activa esta formada por homodímeros de la zona c-terminal, aunque existen casos de heterodímeros entre diferentes miembros de la familia, que presentan actividades biológicas distintas a aquellas de sus homodímeros correspondientes.

Analizando la secuencia aminoacídica de la miostatina bovina usando el método de von Heijne (1987) se encontró un residuo de 23 aminoácidos que constituye el péptido señal para la secreción. En el péptido aminoterminal latente se encuentran dos posibles lugares de N-glicosilación, así como 4 residuos cisteína, que probablemente puedan estar involucrados en la formación de puentes disulfuro, bien para la dimerización o bien para la unión a una proteína latente. El dominio carboxiterminal bioactivo de la miostatina contiene 112 aminoácidos, generados a partir del procesado proteolítico del fragmento N-terminal. Dentro de este fragmento se encuentran 9 residuos de cisteína y uno de glicina en posiciones conservadas. De ellos, 8 residuos de cisteína están involucrados en la formación de puentes disulfuro intramoleculares, y solo uno en puentes disulfuro entre moléculas. En la Figura 40 se muestra la estructura típica de un miembro de la familia TGF- $\beta$ .

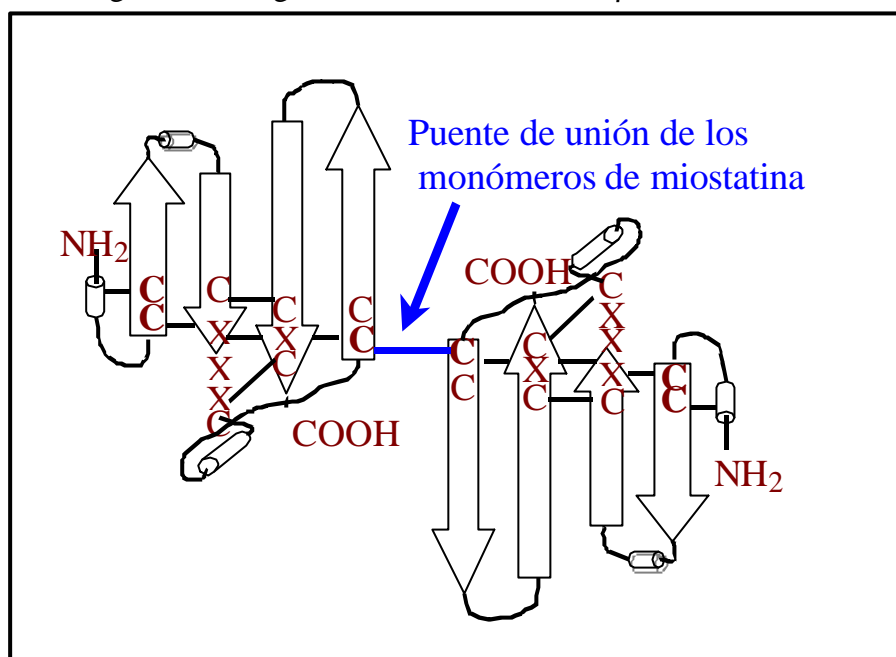
El gen de la miostatina bovina consta de 3 exones interrumpidos por dos intrones. El monómero bioactivo de la miostatina está totalmente codificado por el exón 3. No hay intrones que interrumpan la secuencia no traducida en 3'. Este esquema de organización es el mismo en el resto de los miembros de la familia TGF- $\beta$  (Gonzalez-Cadavid y col., 1998; Guron y col, 1995).

- Figura 40. Estructura molecular de la miostatina bovina.

- a) Estructura lineal de la proteína.



- b) Estructura del homodímero de miostatina, supuesta por analogía a los integrantes de la familia TGF- $\beta$ .

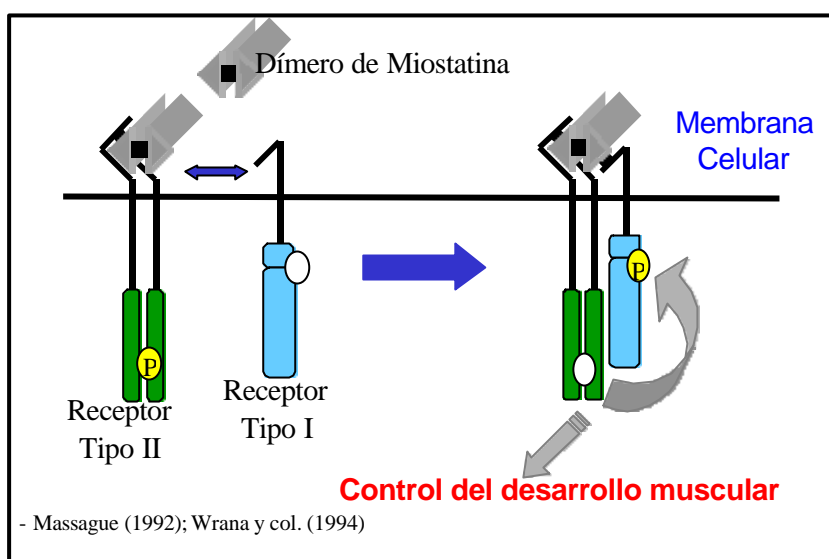


### 3.2 Mecanismo de acción biológica.

La miostatina se expresa fundamentalmente en músculo esquelético y se secreta en plasma (Gonzalez-Cadavid y col., 1998). Probablemente actúa como un *chalon*, es decir como una señal secretada por la célula muscular esquelética que provoca un *feed-back* negativo, directo (señal paracrina) o indirecto en las células precursoras musculares, ejerciendo una regulación de la masa muscular (Slack, 1997). Existen evidencias preliminares de un incremento de las concentraciones de ARN mensajero de miostatina cuando se induce la diferenciación miogénica en mioblastos C2C12 y de un efecto inhibitorio de la miostatina recombinante en la proliferación de mioblastos C2C12 no diferenciados. Esto podría sustentar la hipótesis del efecto paracrina de *feed-back* negativo de la miostatina (D. Poncelet, comunicación personal). Además se ha descrito un efecto de inhibición de la proliferación de mioblastos después de ser incubados con miostatina exógena (Thomas y col., 2000).

El mecanismo de acción general de la miostatina podría ser el siguiente: El dímero de miostatina (molécula bioactiva) se une a un receptor de membrana de la familia de los receptores transmembrana serina-treonina quinasa que está constitutivamente fosforilado (receptor tipo II). Esta unión receptor-ligando (dímero de miostatina) recluta a otro receptor de membrana, el receptor tipo I. La unión al complejo del receptor tipo I, que está constitutivamente defosforilado, provoca la defosforilación del receptor tipo II, fosforilación y consecuente activación del tipo I, desencadenando la respuesta intracelular, encaminada a controlar el desarrollo muscular (Massague, 1992; Wrana y col., 1994). En la Figura 41 se muestra este mecanismo de acción general.

- Figura 41. Posible mecanismo de acción de la miostatina por analogía a los integrantes de la familia TGF- $\beta$ .



#### 4 Estudio de las secuencias repetidas.

El genoma bovino, como el del resto de vertebrados, contiene una gran cantidad de secuencias de ADN altamente repetitivo. Estas secuencias repetitivas se organizan en dos grandes grupos:

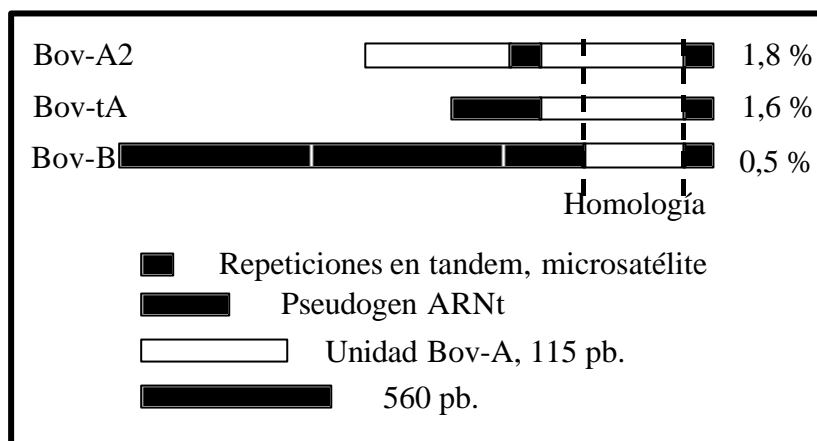
- ADN repetido en tandem: Estas familias están definidas por bloques de ADN repetido en tandem. Dependiendo del tamaño de las unidades de las repeticiones, este ADN no codificante altamente repetitivo se puede agrupar en tres sub-grupos: ADN satélite, minisatélites y microsatélites.

- ADN repetido y disperso: Las unidades de repetición no se encuentran formando grupos, y están dispersas en numerosas localizaciones en el genoma. La mayoría de ellos son capaces de llevar a cabo una retrotransposición, es decir una transposición a través de un ARN intermediario. Estas secuencias dispersas se clasifican según su tamaño en largas o LINE (*Long Interspersed Repetitive Elements*) (Di Nocera y Sakaki, 1990; Martin, 1991), y cortas o SINE (*Short Interspersed Repetitive Elements*) (Okada, 1991).

En bovino se han caracterizado tres familias principales de elementos SINE (Lenstra y col., 1993; Jobse y col., 1995), denominados siguiendo la nomenclatura propuesta por Lenstra, Bov-A2, Bov-tA, y Bov-B (Figura 42). Existe un segmento de 115 pb., conocido como elemento A, que si se encuentra acoplado a otro elemento A orientado en la misma dirección y separados por un espaciado, constituye el denominado Bov-A2 (secuencia consenso X64126). Cuando este elemento A se encuentra asociado a un pseudogen tipo ARNt, forma el llamado Bov-tA (X64124). El tercer tipo, Bov-B, está formado por una secuencia de unos 560 pb. parcialmente homóloga al elemento A (X64125, números de

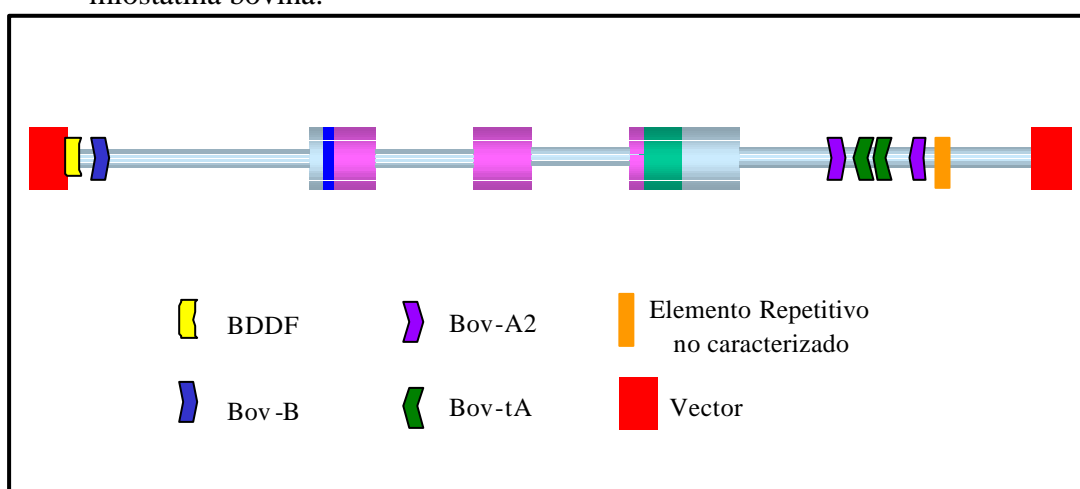
acceso en Lenstra y col., 1993). Estos elementos ocupan aproximadamente 1,8 % Bov-A2, 1,6 % Bov-tA y 0,5 % Bov-A2 en el genoma bovino(Lenstra y col., 1993).

- Figura 42. Secuencias SINE bovinas.



El análisis de la secuencia completa de la miostatina bovina, en busca de homologías con las secuencias almacenadas en *Genbank* y EMBL reveló la presencia de varios elementos repetitivos dispersos bovinos. Para describirlos se sigue la nomenclatura propuesta por Lenstra y col. (1993). Casi todos ellos conservaban una gran homología con las secuencias consenso de los elementos descritos, lo que permitió caracterizarlos y orientarlos fácilmente. En la región 3' del gen se encontraron dos elementos SINE Bov-A2 y otros dos elementos SINE Bov-tA. Además se encontró otro elemento repetitivo, que se encuentra también en ovino, no caracterizado. En la región 5' del gen se encontró un elemento SINE Bov-B, y un segmento de un retroelemento BDDF (Szemraj y col., 1995). Se encuentran, por tanto, 2 elementos Bov-A2 y Bovt-A, que son los más abundantes, y uno del tipo Bov-B, el menos frecuente de los tres. No se encontraron más homologías, excepto con las secuencias de miostatina de otras especies. La localización de estos elementos así como su orientación puede verse en la Figura 43.

- Figura 43. Localización de los elementos repetitivos en la secuencia de la miostatina bovina.



## 5 Comparación de las secuencias bovina y murina del gen de la miostatina.

Como se ha visto en el capítulo 3, cuando se comparan las dos secuencias se observa una línea diagonal que indica la similitud de las dos secuencias, que aparece a partir de la base 1.300 de la secuencia bovina, únicamente interrumpida después entre las bases 12.330 y 15.850.

El inicio de la similitud coincide con los límites de un elemento repetitivo Bov-B en la secuencia bovina. Además en 5' a este elemento SINE, se encuentra en la secuencia bovina otro fragmento de un retroelemento BDDF. Esto parece indicar que el motivo de esta falta de homología entre las secuencias sea la presencia de una inserción, que incluye estos 2 elementos repetitivos, en la secuencia bovina.

El otro lugar de ausencia de homología entre las dos secuencias, probablemente se trate de otro proceso de inserción en la secuencia bovina ocurrido durante la divergencia entre las especies bovina y murina. En este fragmento de secuencia se encuentra en la secuencia bovina la presencia de varios elemento repetitivos SINE, lo que favorece la hipótesis de un proceso de inserción mediado por retroelementos de un segmento cromosómico de unos 3.500 pares de bases en la secuencia bovina.

## 6- Regulación de la transcripción del gen de la miostatina.

La miostatina se detecta en embriones de ratón por primera vez a la edad de 9,5 días *post-coitum* en el miotomo de los somitos. En estados posteriores del desarrollo el ARN mensajero de la miostatina se detecta en un amplio grupo de músculos en desarrollo (McPherron y col., 1997). En cerdos se detectan los mayores niveles de ARN mensajero de miostatina durante los últimos estadios de la gestación, y disminuyen después del nacimiento. En cualquier caso, la miostatina se detecta casi exclusivamente en músculo esquelético adulto durante toda la vida, aunque también se detecta en tejido adiposo. Además se ha puesto de manifiesto la presencia de ARN mensajero de miostatina después del nacimiento en músculo cardíaco y glándula mamaria (Ji y col., 1998). En músculo esquelético en fase de regeneración el ARN mensajero de miostatina se detecta en las células mononucleadas de la periferia de las fibras musculares.

Se han identificado varias regiones conservadas (Capítulo 3) en las secuencias obtenidas de los genes de la miostatina bovina (Capítulo 2) y de ratón (L. Grobet y V. Marot, comunicación personal), y se propone que el motivo por el cual esas regiones están conservadas es porque tienen alguna función en la regulación de la transcripción del gen. De las regiones identificadas fuera del ADN codificante, una de ellas, localizada inmediatamente en 5' del inicio de la transcripción, corresponde al promotor mínimo. En ella se encontraron una caja TATA-box y una caja CAAT-box (ver Figura 39).

Los módulos reguladores específicos musculares normalmente se localizan en la región 5' del promotor, en el primer intrón y ocasionalmente en la región 3' del último

exón (Wasserman y Fickett, 1998). Las regiones identificadas en el gen de la miostatina se localizan en 5' del promotor y en el tercio posterior del primer intrón, con lo que su localización apunta a que tengan una función en la regulación del gen. Sin embargo los análisis de regresión logística para discriminar módulos reguladores de la expresión específica muscular no fueron capaces de identificar, dentro de las otras zonas en 5' y primer intrón, ninguna región con un valor significativo como para ser considerada un módulo funcional (Wasserman, comunicación personal).

En definitiva el análisis de regresión logística aplicado a las regiones conservadas entre bovino y ratón no ha evidenciado la existencia de ningún elemento regulador músculo-específico. Sin embargo esto no está en contradicción con las características de expresión músculo específica de la miostatina, sino que puede ser la evidencia de la existencia de módulos reguladores en *cis* aún no identificados, o que los elementos típicos de la expresión muscular se encuentran localizados fuera de las regiones analizadas. En cualquier caso el análisis de *footprints* filogenéticos ha puesto en evidencia unas regiones candidatas, que más adelante podrían ser caracterizadas por análisis funcionales *in vivo*, para demostrar su función biológica de regulación de la expresión del gen de la miostatina. Fundamentalmente son dos los tipos de estrategias que se pueden llevar a cabo:

1. Eliminar las regiones con secuencias homólogas y observar como se modifica, si se modifica, la expresión del gen.
2. Analizar la actividad de esas regiones en modelos transgénicos.

Existen individuos de fenotipo culón, donde el estudio de la secuencia codificante de la miostatina no ha revelado la existencia de una mutación que explique su fenotipo. En ovejas de raza Texel, de fenotipo culón, el análisis de ligamiento indica la existencia de un locus relacionado con la hipertrofia muscular de estos animales en desequilibrio de ligamiento con los marcadores del cromosoma 2 bovino, indicando la posibilidad de que se trate de la miostatina. Sin embargo en las secuencias codificantes del gen de la miostatina ovina de individuos hipermusculados no se encontró evidencia alguna que pueda asociarse al fenotipo (Marcq y col. 1998).

Parece complicado que exista otro gen relacionado con la hipertrofia muscular situado en la misma región. En consecuencia, la búsqueda de la explicación del fenotipo de estos animales podría estar en la presencia de mutaciones fuera de las secuencias codificantes. Regiones importantes fuera de las codificantes son por ejemplo las regiones implicadas en el proceso de maduración por corte y empalme (*splicing*), es decir en las uniones intron/exon y las dianas de procesamiento. Miranda y col., 2001, han revisado estas regiones mediante protocolos de SSCP, y no han encontrado ninguna mutación que pueda ser relacionada con la presencia del fenotipo culón en estos animales.

Otras regiones importantes donde buscar la explicación a los fenotipos anteriores podrían ser los lugares de poliadenilación, y las regiones identificadas como importantes en la regulación del gen de la miostatina en el transcurso de este trabajo.



## Conclusiones.

---

En esta memoria se han revisado los procedimientos que pueden utilizarse para la identificación de genes responsables de caracteres monogénicos, utilizando como ejemplo el clonado posicional del gen responsable de la hipertrofia muscular bovina. Además, se ha obtenido una secuencia de la miostatina bovina de 17.417 pb. que se ha utilizado en un protocolo de secuenciación comparativa con una secuencia de 15.000 pb. de ratón para intentar identificar las posibles regiones reguladoras de la transcripción del gen de la miostatina.

Los objetivos principales de este trabajo eran el clonado posicional del gen responsable de la hipertrofia muscular bovina y una vez identificado el gen responsable avanzar en el conocimiento de su secuencia y de la regulación de su transcripción.

Las conclusiones de este trabajo son:

- 1) El *locus mh*, localizado en la región BTA2q12, se sitúa en un intervalo cromosómico de 3 cM. flanqueado por los marcadores microsatélite Bulge18 y Bulge20
- 2) Se han identificado en el transcrito de la miostatina bovina dos posibles lugares de inicio y tres posibles lugares de final de la transcripción.
- 3) El promotor del gen de la miostatina es un promotor dirigido por una TATA-box, situada a -21 pb del punto de inicio<sup>2</sup> de la transcripción. Además contiene una CAAT-box a -65 pb del punto de inicio<sup>2</sup>.
- 4) En el proceso de divergencia entre las especies bovina y murina se han producido dos eventos de inserción mediada por retroelementos en la secuencia bovina.
- 5) Se han identificado 3 regiones, dos de ellas en la región 5' y una en el intrón 1, como posibles regiones importantes implicadas en la regulación de la transcripción del gen de la miostatina.
- 6) No se ha evidenciado la existencia de ningún elemento regulador músculo-específico en ninguna de las tres regiones identificadas.

## Referencias bibliográficas.

- Abbott C, Povey S. (1991). Development of human chromosome-specific PCR primers for characterization of somatic cell hybrids. *Genomics*, **9**(1): 73-7.
- Anderson (1998). Identification and cloning of trait genes. In *Animal Breeding: Technology for the 21<sup>st</sup> century*, J. Clark, ed. (Amsterdam: Harwood Academic Publishers), pp 103-118.
- Andersson B., Lu J., Shen Y., Wentland M.A., Gibbs R.A. (1997). Simultaneous shotgun sequencing of multiple cDNA clones. *DNA Sequencing*, **7**: 63-70.
- Antoniou E., Grosz M.(1999). PCR based detection of bovine myostatin Q204X mutation. *Animal Genetics*, **30**: 231-232.
- Archibald (1998). Comparative genome mapping-the livestock perspective. In *Animal Breeding: Technology for the 21<sup>st</sup> century*, J. Clark, ed. (Amsterdam: Harwood Academic Publishers), pp 137-164.
- Arnold C. y Hodgson I.J. (1991). Vectorette PCR: a novel approach to genomic walking. *PCR Methods Appl.* **1**: 39-42.
- Arthur P.F., Makarechian M.Y., Price M.A. (1988). Incidence of distocia and perinatal calf mortality resulting from reciprocal crossing of double-muscléd and normal cattle. *Canadian Veterinary Journal*, **29**:163-167.
- Arthur P.F. (1995). Double muscling in cattle: a review. *Aust. J. Agric. Res.*, **46**: 1493-1515.
- Bahary N., Pachter J.E., Felman R., Leibel R.L., Albright K., Cram S. Y Friedman J.M.(1992). Molecular mapping of mouse chromosome 4 and 6: use of a flow-sorted robertsonian chromosome. *Genomics*, **13**: 761-7.
- Bankier A.T., Weston K.M., Barrell B.G. (1987). Random cloning and sequencing by the M13/dideoxynucleotide chain termination method. *Methods Enzymology*, **155**: 51-93.
- Baron B., Metezeau P., Kelly F., Bernheim A., Berger R., Guenet J.L. y Goldberg M.E. (1984). Flow cytometry isolation and improved visualization of sorted mouse chromosomes. Purification chromosomes X and ISO-1 from cell lines with Robertsonian translocations. *Exp. Cell. Res.*, **152**: 220-30.
- Baron B., Metezeau P., Kiefer-Gachelin H. y Goldberg M.E.(1990). Construction and characterization of a DNA library from mouse chromosomes 19 purified by flow cytometry. *Biol. Cell.*, **69**: 1-8.
- Bellanné-Chantelot C., Lacroix B., Ougen P., Billaut A., Beaufils S., Bertrand S., Georges I., Gilbert F., Gros I., Lucotte G., Susini L., Codani J.-J., Gesnouin P., Pook S., Vaysseix G., Lu-Kuo J., Ried T., Ward D., Chumakov I., Le Paslier D., Barillot E., Cohen D. (1992). Mapping the whole human genome by fingerprinting Yeast Artificial Chromosomes. *Cell*, **70**: 1059-1068.
- Boccard R. Y Dumont B.L.(1974). Consequences de l'hipertrophy musculaire héréditaire des bovins sur la musculature. *Ann. Génét. Sél. Anim.*, **6**: 177-186.
- Boccard R. Y Monin G. (1973). Glycogénolyse *post mortem* et caractéristiques de qualité de la viande des bovins culards. 7p.

- Boccard R. (1982). Relationship between muscle hypertrophy and the composition of skeletal muscles. En: "Muscle Hypertrophy of Genetic Origin and its Use to Improve Beef Production". (Eds J.W.B. King and F. Ménéssier) pags. 148-162. (Martinus Nijhoff Publishers, The Hague).
- Boyajeau D., Ménéssier F., Valls Ortiz J.M. (1971). Document de travail sur le caractère culard chez les bovins. 38p.
- Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D., Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D., Kerlavage A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reich C.I., Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A., Scott J.L., Geoghagen N.S.M., Venter J.C. (1996). Complete genome sequence of the methanogenic archaeon *Methanococcus jannaschii*. *Science*, **273**: 1058-1073.
- Cañón J., Gutierrez J.P., Dunner S., García-Atance P., Goyache F., Villa A., Alonso L. (1996). Principales aspectos del programa de mejora genética de la raza Asturiana de los Valles. ITEA 92:75-91
- Capuccio I., Marchitelli C., Serracchioli A., Nardone A., Filippini F., Ajmone-Marsan P., Valentini A. (1998). A G-T transversion introduces a stop codon at the *mh* locus in hypertrophic Marchigiana beef subjects. *Animal Genetics*, **29** (suppl. 1): 51.
- Carter N.P., Ferguson-Smith M.A., Perryman M.T., Telenius H., Pelemear A.H., Leversha M.A., Glancy M.T., Wood S.L., Cook K., Dyson H.M. y col. (1992). Reverse chromosome painting: a method for the rapid analysis of aberrant chromosomes in clinical cytogenetics. *J. Med. Genet.*, **29**: 299-307.
- Charlier C., Coppieters W., Farnir F., Grobet L., Leroy P.L., Michaux C., Mni M., Schwers A., Vanmanshoven P., Hanset R., Georges M. (1995). The *mh* gene causing double-muscling in cattle maps to bovine Chromosome 2. *Mammalian Genome*, **6**: 788-792.
- Charlier C., Denys B., Belanche J.I., Coppieters W., Grobet L., Mni M., Womack J., Hanset R., Georges M. (1996a). Microsatellite mapping of a major determinant of White Heifer Disease: the bovine *roan* locus. *Mammalian Genome*, **7**: 138-142.
- Charlier C., Farnir F., Berzi P., Vanmanshoven P., Brouwers B., Vromans H., Georges M. (1996b). Identity-by-descent mapping of recessive traits in livestock: application to map the bovine syndactyl *locus* to chromosome 15. *Genome Research*, **6**: 580-589.
- Charlier C. (1999). Application des méthodes de cartographie génétique à l'étude des caractères monogéniques chez le bovin. Thèse présentée en vue de l'obtention du grade de Docteur en Sciences Vétérinaires. Université de Liège.
- Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J., Rutter, W.J. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry*, **18**: 5294-5299.
- Collins F.S. (1995). Positional cloning moves from perditional to traditional. *Nature Genetics*, **9**: 347-350.
- Copeland NG, Jenkins NA, Gilbert DJ, Eppig JT, Maltais LJ, Miller JC, Dietrich WF, Weaver A, Lincoln SE, Steen RG, y col. (1993). A genetic linkage map of the mouse: current applications and future prospects. *Science*, **262**: 57-66.
- Cox D.R., Burmeister M., Price E.R., Kim S., Myers R.M. (1990). Radiation hybrid mapping: a somatic cell genetic method for constructing high-resolution maps of mammalian chromosomes. *Science*, **250**: 245-250.
- Cram L.S. (1990) Flow cytogenetics and chromosome sorting. *Hum. Cell*, **3**: 99-106.
- Crete N., Delabar J.M., Sinet P.M., Creau-Goldberg N. (1991). Accurate evaluation of the sizes of DNA fragments (from 30 to 4700 kb.) in pulse field gel electrophoresis. *Biotechniques*, **11**: 711-716.

- Culley G. (1807) Observation on livestock, 4<sup>th</sup> ed. London, G. Woodfall.
- Di Nocera P.P. y Sakaki Y. (1990). LINEs: a superfamily of retrotransposable ubiquitous elements. *Trends in Genetics*, **6**:29-30.
- Dunner S., Charlier C., Farnir F., Brouwers B., Cañón J., Georges M. (1997). Towards interbreed IBD mapping of the *mh* locus: double-muscling in the *Asturiana de los Valles* breed involves the same locus as in the *Belgian Blue* cattle breed. *Mammalian Genome*, **8**: 430-435.
- Dutrillaux B. (1979). Chromosomal evolution in primates: tentative phylogeny from *Microcebus murinus* (Prosimian) to man. *Human Genetics*, **48**: 251-314.
- Dutrillaux B., Viegas-Péquignot E., Couturier J. (1980). Très grande homologie de marquage chromosomique entre le lapin (*Oryctolagus cuniculus*) et les primates, dont l'homme. *Annals Génétiques*, **23**: 22-25.
- Ewing, B., Hillier, L., Wendl, M.C., Green, P. (1998). Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. *Genome Research*, **8**(3):175-185.
- Ewing, B. y Green, P. (1998). Base-calling of automated sequencer traces using Phred. II. Error probabilities. *Genome Research*, **8**(3):186-194.
- Fahrenkrug S.C., Casas E., Keele J.W., Smith T.P. (1999). Technical note:direct genotyping of the double-muscling locus (mh) in Piedmontese and Belgian Blue Cattle by Fluorescent PCR. *J. Anim. Sci.*, **77**: 2028-2030.
- Fernández C., Georges M., Grobet L., Cañón J., Dunner S. (2002). Generación en células *ES* murinas de un alelo del gen de la miostatina susceptible de activación bajo la dependencia de CRE. En: *XI Reunión Nacional de Mejora Genética Animal*, Pamplona 2002.
- Fickett J.W. y Wasserman W.W. (2000). Discovery and modeling of transcriptional regulatory regions. *Current Opinion in Biotechnology*, **11**: 19-24.
- Fisher S.R., Beever J.E., Lewing H.A. (1996). Genetic mapping of Col3A1 to bovine chromosome 2. *Mammalian Genome*, **8**: 76-77.
- Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F., Kerlavage A.R., Bult C.J., Tomb J.F., Dougherty B.A., Merrick J.M., y col. (1995). Whole-genome random sequencing and assembly of *Haemophilus influenzae*. *Science*, **269**: 496-512.
- Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A., Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M., y col. (1995). The minimal gene complement of *Mycoplasma genitalium*. *Science*, **279**: 397-403.
- García Fierro B.F. (1972). El ganado vacuno en Asturias. El carácter “ Anca de potro”, “Grupa doble” o “Culón”. *Ganadería*, **345**: 117-128.
- Gardiner K., Laas W., Patterson D. (1986). Fractionation of large mammalian DNA restriction fragmentes using vertical pulsed-field gradient gel electrophoresis. *Somat Cell Molecular Genetics*, **12**: 185-195.
- Georges M., Lathrop M., Gilbert M., Marcotte A., Schwers A., Swillens S., Vassart G., Hanset R. (1990). On the use of DNA fingerprints for linkage studies in cattle. *Genomics*, **6**: 461-474.
- Georges M., Lathrop M., Dietz A.B., Lefort A., Libert F., Mishra A., Nielsen D., Sargeant L.S., Steele M.R., Zhao X., Leipold H., Womack J.E. (1993a). Microsatellite mapping of a gene causing weaver disease in cattle will allow the study of an associated QTL. *Proc. Natl. Acad. Sci. U.S.A.*, **90**: 1058-1062.

- Georges M., Drinkwater R., Lefort A., Libert F., King T., Mishra A., Nielsen D., Sargeant L.S., Sorensen A., Steele M.R., Zhao X., Womack J.E., Hetzel J. (1993b). Microsatellite mapping of a gene affecting horn development in *Bos taurus*. *Nature Genetics*, **4**: 206-210.
- Georges (1998). Mapping genes underlying production traits in livestock. In *Animal breeding: Technology for the 21<sup>st</sup> century*, J. Clark, ed. (Amsterdam: Harwood Academic Publishers), pp 77-101.
- Gerrard D.E., Thrasher K.H., Grant K.L., Lemenager R.P. y Judge M.D. (1991). Serum-induced myoblast proliferation and gene expression during development of double muscled and normal cattle. *Journal of Animal Science*, **69** (supl. 1): 317.
- Geurts J.M.W., Schoenmakers E.F.P.M., Mols R., Van de Ven W.J.M. (1994). Improved procedure for rapid isolation and sequencing of DNA insert termini in Yeast Artificial Chromosomes. *Methods in Molecular and Cellular Biology*, **4**: 257-265.
- Gonzalez-Cadavid N.F., Taylor W.E., Yarasheski K., Sinha-Hikim I., Ma K., Ezzat S., Shen R., Lalani R., Asa S., Mamita M., Nair G., Arver S., Bhasin S. (1998). Organization of the human myostatin gene and expression in healthy men and HIV-infected men with muscle wasting. *Proc. Natl. Acad. Sci. USA*, **95**: 14938-14943.
- Gordon, D., Abajian, C., Green, P. (1998). Consed: a graphical tool for sequence finishing. *Genome Research* 8(3): 195-202.
- Gray J.W. y Langlois R.G. (1986). Chromosome classification and purification using flow cytometry and sorting. *Ann. Rev. Biophys. Chem*, **15**: 195-235.
- Grobet L., Royo L.J., Poncelet D., Pirottin D., Brouwers B., Riquet J., Schoberlein A., Dunner S., Menissier F., Massabanda J., Fries R., Hanset R., Georges M. (1997). A deletion in the bovine myostatin gene caused the double-muscled phenotype in cattle. *Nature Genetics*, **17**: 71-74.
- Grobet L., Poncelet D., Royo L.J., Brouwers B., Pirottin D., Michaux C., Menissier F., Zanotti M., Dunner S., Georges M. (1998). Molecular definition of an allelic series of mutations disrupting the myostatin function and causing double muscled in cattle. *Mammalian Genome*, **9**: 210-213.
- Grossman PD, Bloch W, Brinson E, Chang CC, Eggerding FA, Fung S, Iovannisci DM, Woo S, Winn-Deen ES (1998). High-density multiplex detection of nucleic acid sequences: oligonucleotide ligation assay and sequence-coded separation. *Nucleic Acids Res* 22(21):4527-4534.
- Grunwald D., Geffrotin C., Chardon P., Frelat G. y Vaiman M. (1986). Swine chromosomes: flow sorting and spot blot hybridization. *Cytometry*, **7**: 582-8.
- Guérin G., Duval-Iflah Y., Bonneau M., Bertaud M., Guillaume P., Olivier L. (1993). Evidence for linkage between K88ab, K88ac intestinal receptors to *Escherichia coli* and transferrin loci in pigs. *Animal Genetics*, **24**, 393-396.
- Guron C., Sudarshan C., Raghow R. (1995). Molecular organization of the gene encoding murine transforming growth factor beta 1. *Gene*, **165**:325-6.
- Hanset R. y Michaux C. (1982). Creatine and creatinine levels in plasma red cells and muscles as characteristics of double-muscled cattle. En: *Muscle hypertrophy of genetic origin and its use to improve beef production*. Eds. J.W.B. King and F. Ménissier, pp. 237-256. (Martinus Nijhoff Publishers: The Hague).
- Hanset R. y Michaux C. (1985a). On the genetic determinism of muscular hypertrophy in the Belgian White and Blue cattle breed. I. Experimental data. *Génét. Sél. Evol.*, **17**, 359-368.
- Hanset R. y Michaux C. (1985b). On the genetic determinism of muscular hypertrophy in the Belgian White and Blue cattle breed. II. Population data. *Génét. Sél. Evol.*, **17**, 369-386.

- Hardison R.C., Oeltjen J., Miller W. (1997). Long human-mouse sequence alignments reveal novel regulatory elements: a reason to sequence the mouse genome. *Genome Research*, **7**: 959-966.
- Holmes J.H. y Robinson D.W. (1970). Hereditary muscular hypertrophy in the bovine: metabolic response to nutritional stress. *J Anim Sci.*, **31**: 776-80
- Holmes J.H.G. y Ashmore C.R. (1972). A histochemical study of development of muscle fibre type and size in normal and "double-muscled" cattle. *Growth*, **36**: 351-372.
- Holmes J.H.G. y Robinson D.W. (1970). Hereditary muscular hypertrophy in the bovin: methabolic response to nutritional stress. *J. Anim. Sci.*, **31**: 776-780.
- Ji S, Losinski RL, Cornelius SG, Frank GR, Willis GM, Gerrard DE, Depreux FF, Spurlock ME. (1998). Myostatin expression in porcine tissues: tissue specificity and developmental and postnatal regulation. *Am J Physiol.*, **275**: 265-73.
- Jobse C., Buntjer J.B., Haagsma N., Breukelman H.J., Beintema J.J., Lenstra J.A. (1995). Evolution and recombination of bovine DNA repeats. *Journal of Molecular Evolution*, **41**: 277-283.
- Kaczorowski T. y Szybalski W. (1998). Genomic DNA sequencing by SPEL-6 primer walking using hexamer ligation. *Gene*, **223**: 83-91.
- Kambadur R., Sharma M., Smith T.P.L., Bass J.J. (1997). Mutations in myostatin (GDF-8) in double-muscled Belgian Blue cattle. *Genome Research*, **7**: 910-916.
- Kappes S.M., Keele J.W. Stone R.T., McGraw R.A., Sonstegard T.S., Smith T.P., Lopez-Corrales N.L., Beattie C.W. (1997). A second-generation linkage map of the bovine genome. *Genome Research*, **7**(3): 235-249.
- Karim L., Coppieters W., Grobet L., Valentini A., Georges M. (2000). Convenient genotyping of six myostatin mutations causing double-muscling in cattle using a multiplex oligonucleotide ligation assay. *Animal Genetics*, **31**: 396-399
- Kaukinen J. y Varvio S.-L. (1992). Artiodactyl retroposons: association with microsatellites and use in SINEmorph detection by PCR. *Nucleic Acids Research*, **20**: 2955-2958.
- Kehrli M.E. Jr., Schmalstieg F.C., Anderson D.C., Van Der Maaten M.J., Hughes B.J., Ackermann M.R., Wilhelmsen C.L., Brown G.B., Stevens M.G., Whetstone C.A., (1990). Molecular definition of the bovine granulocytopeny syndrome: identification of deficiency of the Mac-1 (CD11b/CD18) glycoprotein. *American Journal of Veterinary Research*, **51**: 1826-1836.
- Keller E.B. y Noon W.A. (1984) Intron splicing: a conserved internal signal in introns of animal pre-mRNAs. *Proc Natl Acad Sci U S A* ;**81**: 7417-20.
- Kemp S.J., Hishida O., Wambugu J., Rink A., Longeri M.L., Ma R.Z., Da Y., Lewin H.A., Barendse W., Teale A.J. (1995). A panel of polymorphic bovine, ovine, and caprine microsatellite markers. *Animal Genetics*, **26**: 299-306.
- Kidwell J.F., Vernon E.H., Crown R.M., Sigletary C.B. (1952). Muscular hypertrophy in cattle. *J. Hered.*, **43**: 62-68.
- Klungland H., Vage D.I., Raya L.G., Adalsteinsson S., Lien S. (1995). The role of the melanocyte stimulating hormone (MSH) receptor in bovine coat color determination. *Mammalian Genome*, **6**: 636-639.
- Kostia S., Vilkkij J., Pirinen M., Womack J.E., Barendse W., Varvio S.-L. (1997). SINE targeting of bovine microsatellites from bovine/rodent hybrid cell lines. *Mammalian Genome*, **8**: 365-367.
- Kozak M. (1996). Interpreting cDNA sequences: some insights from studies on translation. *Mammalian Genome*, **7**: 563-574.

- Kronacher C. (1934). Genetik und Tierzucht. En Baur E. y Hartman M., *Handbuch der Verebungswissenschaft*, Gebrüder Borntrager, Berlin, **3**: 139 pp.
- Lanneluc I., Drinkwater R.D., Elsen J.M., Hetzel J., Nguyen T., Piper L.P., Thimonier J., Harrison B., Gellin J. (1994). Genetic markers for the Booroola fecundity (Fec) gen in sheep. *Mammalian Genome*, **5**: 26-33.
- Lee T.I. y Young R.A. (2000). Transcription of eukariotic protein-coding genes. *Annu. Rev. Genet.*, **34**: 77-137.
- Lenstra J.A., van Boxtel J.A.F., Zwaagstra K.A., Schwerin M. (1993). Short interspersed nuclear element (SINE) sequences of the *Bovidae*. *Animal Genetics*, **24**: 33-39.
- Libert F., Lefort A., Okimoto R., Georges M. (1993). Construction of a bovine genomic library of large yeast artificial chromosome clones. *Genomics*, **18**: 270-276.
- Lisitsyn N.A., Lisitsyn N., Wigler M. (1993). Cloning the differences between two complex genomes. *Science*, **259**: 946-951.
- Lüdecke H.-F., Senger G., Claussen u., Horsthemke B.(1989). Cloning defined regions of the human genome by microdissection of banded chromosomes and enzymatic amplification. *Nature*, **338**: 348-350.
- Lyons L.A., Laughlin T.F., Copeland N.G., Jenkins N.A., Womack J.E., O'Brien S.J. (1997). Comparative anchor tagged sequences (CATS) for integrative mapping of mammalian genomes. *Nature Genetics*, **15**: 47-56.
- Ma R.Z., Beever J.E., Da Y., Green C.A., Russ I., Park C., Heyen D.W., Everts R.E., Fisher S.R., Overton K.M., Teale A.J., Kemp S.J., Hines H.C., Guerin G., Lewin H.A. (1996). A male linkage map of the cattle (*Bos taurus*) genome. *Journal of Heredity*, **87**(4): 261-271.
- Magliano A. (1933). Indagini circa il comportamento genetico della cosiddetta "gropa di cavallo" nei bovini di razza Piemontese. *Clinica Vet.*, **11**: 124-139.
- Mairal A. (1994). Caracterisation d'un panel d'hybrides somatiques interspecifiques par cytogenetique moleculaire. Memoire our el Diplome d'etudes Superieures. Université Paul Sabatier-Toulouse III.
- Maizel y Lenk (1981). Enhanced graphic matrix analysis of nucleic acid and protein sequences. *Proc Natl Acad Sci U S A.*, **78**:7665-7669.
- Marcq F., Elsen J.-M., El Barkouki S., Bouix J., Eychenne F., Grobet L., Karim L., Laville E., Nezer C., Royo L.J., Said T., Bibe B., Le Roy P.L., Georges M. (1998). Investigatin the role of myostatin in the determinism of double muscling characterizing Belgian Texel sheep. *Animal Genetics*, **29** (Suppl. 1): 52.
- Martin S.L. (1991). LINEs. *Current Opinion in Genetics and Development*, **1**: 505-508.
- Maruyama I.N., Rakow T.L., Maruyama H.I.(1995). CRACE: a simple method for the identification of the 5' end of mRNAs. *Nucleic Acids Research*, **23**: 3796-3797.
- Masabanda J., Pirottin D., Poncelet D., Grobet L., Georges M., Fries R. (1998). A cytogenetic map of the muscular hypertrophy (MH) region on bovine chromosome 2. *Animal Genetics*, **29**(Suppl. 1): 33.
- Massague J. (1992). Receptors for the TGF- $\beta$  family. *Cell*, **69**: 1067-1070.
- Masoero G. (1982). Creatinine variation in transmission of muscle hypertrophy in Piedmontese cattle. En: *Muscle hipertrophy of genetic origin and its use to improve beef production*. Eds. J.W.B. King and F. Ménissier, pp. 219-235. (Martinus Nijhoff Publishers: The Hague).

- McKellar (1968). The occurrence of muscular hypertrophy in South Devon cattle. *Vet. Rec.*, **72**: 507-510.
- McMurray A.A., Sulston J.E., Quail M.A. (1998). Short-insert libraries as a method of problem solving in genome sequencing. *Genome Research*, **8**: 562-566.
- McPherron A.C., Lee S.-J. (1996). The transforming growth factor  $\beta$  superfamily. En: *Growth factors and citokines in health and disease*. Vol. **1B**: 357-393. 1996 JAI Press Inc.
- McPherron A.C., Lee S.-J. (1997). Double muscling in cattle due to mutations in the myostatin gene. *Proc. Natl. Acad. Sci. USA*, **94**: 12457-12461.
- McPherron A.C., Lawler A.M., Lee S.-J. (1997). Regulation of skeletal muscle mass in mice by a new TGF- $\beta$  superfamily member. *Nature*, **387**: 83-90.
- McPherron A.C., Lawler A.M., Lee S.-J. (1999). Regulation of anterior/posterior patterning of the axial skeleton by growth/differentiation factor 11. *Nat Genet.*; **22**:260-264.
- Ménissier F. (1982). Present state of knowledge about the genetic determination of muscular hypertrophy or the double muscled trait in cattle. In *Current topics in Veterinary Medicine and Animal Science*, vol 16: *Muscle hypertrophy of genetic origin and its use to improve beef production*. Ed. King and Ménissier, *Curr. Top. Vet. Anim. Sci.*, **16**: 387-428.
- Milan D., Riquet J., Yerle M., Goureau A., Schmitz A., Cribiu E.P., Frelat G., Gellin J. (1996). Hologous and heterologous FISH painting with PARM-PCR chromosome-specific probes in mammals. *Mammalian Genome*, **7**: 194-199.
- Miranda M.E., Cañón J., Ménissier F., Hanset R., Williams J., Dunner S. (2001). Identificación de los distintos haplotipos del gen de la miostatina en razas europeas. *ITEA*, **22** Volumen extra, Tomo I: 39.
- Miranda M.E., Amigues Y., Boscher M. Y., Ménissier F., Cortés O., Dunner S. (2002). Simultaneous genotyping to detect myostatin gene polymorphism in beef cattle breeds. *Journal of Animal Breeding and Genetics*, en prensa.
- Nash W.C., O'Brien S.J. (1982). Conserved regions of homologous G-banded chromosomes between orders in mammalian evolution: carnivores and primates. *Proc. Natl. Acad. Sci. USA*, **79**: 6631-6635.
- Naredo M y Bajo F. (1916). El ganado bovino de Asturias. Memorias premiadas en el concurso de 1916 por la asociación General de Ganaderos. Imprenta Alemana, Madrid (España), 184 pp.
- Naylor S.L. (1996). Comparison of human chromosome 2 to the mouse map. *Cytogenetics Cell Genetics*, **73**: 271-272.
- Needleman S.B. y Wunsch C.D. (1979). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.*, **48**: 443-453.
- Nehls M. y Boehm T. (1995). A rapid and efficient alternative to sonication in shot-gun sequencing projects. *Trends in Genetics*, **11**: 39.
- Novakofski J.E. y Kauffman R.G. (1981). Biological detection of heterozygosity for doublemuscling in cattle. II. Thyroid hormone concentrations. *Journal of Animal Science*, **52**: 1437-1441.
- Novakofski J.E., Kauffman R.G., Cassens R.G. (1981). Biological detection of heterozygosity for doublemuscling in cattle. I. Terminal innervation ratio. *Journal of Animal Science*, **52**: 1430-1436.



- Nusgens B.V., Verellen-Dumoulin C., Hermanns-Le T., De Paepe A., Nuytink L., Pierard G.E., Lapiere C.M. (1992). Evidence for a relationship between Ehlers-Danlos type VII C in humans and bovine dermatosparaxis. *Nature Genetics*, **1**: 214-217.
- O'Brien S.J., Womack J.E., Lyons L.A., Moore K.J., Jenkins N.A., Copeland N.G. (1993). Anchor reference loci for comparative genome mapping in mammals. *Nature Genetics*, **3**: 103-112.
- Okada N. (1991). SINEs. *Current Opinion in Genetics and Development*, **1**: 498-504.
- Olson M., Hood L., Cantor C., Botstein D. (1989). A common language for physical mapping of the human genome. *Science*, **245**: 1434-1435.
- Paci C. (1935). Indagini e controlli sulla formazione della sotta razza Albese. *Riv. Zootech.*, **12**: 149-225.
- Pearson W.R. y Lipman D.J (1988). Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA*, **85**: 2444-2448.
- Pirottin D., Poncelet D., Grobet L., Royo, L.J., Brouwers B., Massabanda J., Takeda H., Fries R., Sugimoto Y., Womack J.E., Dunner S., Georges M. (1999). High-resolution, human-bovine comparative mapping based on a closed YAC contig spanning the bovine *mh* locus. *Mammalian genome*, **10**: 289-293.
- Rack K.A., Harris P.C., MacCarthy A.B., Boone R., Raynham H., McKinley M., Fitchett M., Towe C.M., Rudd P., Armour J.A.L. y col. (1993). Characterization of three de novo derivative chromosome 16 by reverse chromosome painting and molecular analysis. *Am. J. Hum. Genet.*, **52**: 987-997.
- Raimondi R. (1957). Studio sui bovin piamontese a "groppe doppie". *Annali Accad. Agric.*, **17**: 471-489.
- Rempel W.E., Lu M., Kandelgy S.E., Kennedy C.F.H., Irvin L.R., Mickelson J.R., Louis C.F. (1993). Relative accuracy of the halotane challenge test and a molecular genetic test in detecting the gene for porcine stress syndrome. *Journal of Animal Science*, **71**: 1395-1399.
- Riquet J. (1996). Cartographie génique du porc: apport de la cytométrie en flux au développement de la carte génétique d'un chromosome, et à la cartographie comparée interspécifique par FISH-Painting. These. Laboratoire de Génétique Cellulaire- INRA.
- Riquet J., Coppieters W., Cambisano N., Arranz J.-J., Berzi P., Davis S.K., Grisart B., Farnir F., Karim L., Mni M., Simon P., Taylor J.F., Vanmanshoven P., Wagenaar D., Womack J.E., Georges M. (1999). Fine-mapping of quantitative trait loci by identity by descent in outbred populations: Application to milk production in dairy cattle. *Proc. Natl. Acad. Sci. USA*, **96**: 9252-9257.
- Sambrook J., Fritsch E.F., Maniatis T. (1989). Molecular cloning: A laboratory manual, 2<sup>nd</sup> edition. Cold Spring Harbor Press, Cold Spring Harbor, N.Y.
- Schmidt W.M. y Mueller M.W. (1999). CapSelect: a highly sensitive method for 5'Cap-dependent enrichment of full-length cDNA in PCR-mediated analysis of mRNAs. *Nucleic Acids Research*, **27**: e31:i-iv.
- Slack J.M.W. (1997). Growth control: action mouse. *Current Biology*, **7**: R467-R469.
- Slate J., Coltman D.W., Goodman S.J., MacLean I., Pemberton J.M., Williams J.L. (1998) Bovine microsatellite loci are highly conserved in red deer (*Cervus elaphus*), sika deer (*Cervus nippon*) and Soay sheep (*Ovis aries*). *Animal Genetics*: **29**(4): 307-315.
- Smith W.H. (1949). The occurrence of the double-muscled character in beef cattle. *M. Sc. Thesis*, Kansas State College, 65p.

- Smith y Waterman (1981). *Advances in Applied Mathematics*, **2**: 482-489.
- Schmitz A., Oustry A., Chaput B., Bahri-Darwich I, Yerle M., Millan D., Frelat G., Crihiu E.P. (1995). The bovine bivariate flow karyotype and peak identification by chromosome painting with PCR-degenerated probes. *Mammalian Genome*, **6**: 415-420.
- Solinas-Toldo S., Lengauer C., Fries R. (1995). Comparative genome map of man and cattle. *Genomics*, **27**: 489-496.
- Sonstegard T.S., Lopez-Corrales N.L., Kappes S.M., Beattie C.W., Smith T.P.L (1997a). Comparative mapping of human chromosome 2 identifies segments of conserved synteny near the bovine mh locus. *Mammalian Genome*, **8**: 751-755.
- Stallings R.L., Ford A.F., Nelson D., Torney D.C., Hildebrand C.E., Moyzis R.K. (1991). Evolution and distribution of (GT)<sub>n</sub> repetitive sequences in mammalian genomes. *Genomics*, **10**:807-815.
- Stone R.T., Pulido J.C., Duyk G.M., Kappes S.M., Keele J.W., Beattie C.W. (1995) A small-insert bovine genomic library highly enriched for microsatellite repeat sequences *Mammalian Genome*. **6**: 714-724.
- Strachan T., y Read A.P. (1999). Human molecular genetics 2. BIOS® Scientific Publishers Ltd.
- Strauss E.C., Kobori J.A., Siu G., Hood L.E. (1986). Specific-primer-directed DNA sequencing. *AnalyticalBiochemistry*, **154**: 353-360.
- Swatland H.J. (1973). Innervation of genetically enlarged muscles from double-musled cattle. *Journal of Animal Science*, **36**: 355-362.
- Szabo G., Dallmann G., Muller G., Patthy L., Soller M., Varga L. (1998). A deletion in the myostatin gene causes the compact (cmpt) hypermuscular mutation in mice. *Mammalian Genome*, **9**(8): 671-672.
- Szemraj J., Plucienniczak G., Jaworski J, Plucienniczak A. (1995). Bovine Alu-like sequences mediate transposition of a new site-specific retroelement. *Gene*, **152**: 261-264.
- Thomas M, Langley B, Berry C, Sharma M, Kirk S, Bass J, Kambadur R.(2000). Myostatin, a negative regulator of muscle growth, functions by inhibiting myoblast proliferation. *J. Biol. Chem.* **275**(51): 40235-43.
- Trask B.J. (1991). Fluorescence *in situ* hybridization: applications in cytogenetics and gene mapping. *Trends in Genetics*, **7**: 149-154.
- Uytterhaegen L., Claeys E., Demeyer D., Lippens M., Fiems L.O., Boucqué C.Y., Van de Voorde G. Y Bastiaens A. (1994). Effects of double-muscling on carcass quality, beef tenderness and myofibrillar protein degradation in Belgian Blue White bulls. *Meat Science*, **38**: 255-267.
- Vaiman D., Mercier D., Moazami-Goudarzi K., Eggen A., Ciampolini R., Lepingue A., Velmala R., Kaukinen J., Varvio S.L., Martin P., Leveziel H., Guerin G. (1994) A set of 99 cattle microsatellites: characterization, synteny mapping, and polymorphism. *Mammalian Genome*, **5**: 288-297.
- Vallejo M., Gutierrez J.P., Cima M., Cañón J., Alonso L., Revuelta J.R., Goyache F. (1993). Características de las canales de las razas bovinas asturianas. III Valoración cuantitativa y predicción de la composición tisular de canales en la raza Asturiana de los Valles. *Archivos de Zootecnia*, **42**: 29-40.
- Van Ommen G.-J. B., Breuning M.H., Raap A.K. (1995). FISH in genome research and molecular diagnostics. *Current Opinion in Genetics and Development*, **5**: 304-308.
- Varga L., Szabo G., Darvasi A., Müller G., Sass M., Soller M. (1997) Inheritance and mapping of *Compact (Cmpt)*, a new mutation causing hypermuscularity in mice. *Genetics*, **147**: 755-764.

- Verhasselt P., Voet M., Volckaert G. (1992). DNA sequencing by a subcloning-walking strategy using a specific and semi-random primer in the polymerase chain reaction. *DNA sequencing*, **2**: 281-287.
- Villa, A. (1999). La raza Asturiana de los Valles. En "El Ganado vacuno del Tronco Castaño", pp.161-182, Imprenta Narcea S.L., Granda-Siero (Asturias).
- Visscher P.M., Van der Beek S., Haley C.S.(1998)Marker assisted Selection. En: *Animal Breeding: Technology for the 21<sup>st</sup> Century*, J Clark, ed. (Amsterdam: Harwood Academic Publishers), pag. 119-136.
- Vissac B., Menissier F. Perreau F.(1973). Etude du caractère culard. VII Croissance et musculature des femelles, déséquilibre morphologique au vêlage. *Ann. Génét. Sél. Anim.*, **5**: 23-38.
- von Heijne G. (1987). SIGPEP: a sequence database for secretory signal peptides. *Protein Seq Data Anal.*, **1**(1): 41-2.
- Warburton D., Gersen D., Yu M-T., Jackson C., Handelin B., Houseman D. (1990). *Genomics*, **6**: 358-366.
- Wasserman W.W. y Fickett J.W. (1998). Identification of regulatory regions wich confer muscle-specific gene expression. *Journal of Molecular Biology*, **278**: 167-181.
- Wasserman W.W, Palumbo M., Thompson W., Fickett J.W., Lawrence C.E. (2000). Human-mouse genome comparisons to locate regulatory sites. *Nature Genetics*, **26**: 225-228.
- Weber A.D. e Ibsen H.L. (1934). The ocurrence of the double-muscled character in pure breed beef cattle. *Proc. Amer. Soc. Anim. Prod.*: 228-232.
- Weber K.L., Bolander M.E., Sarkar G. (1998). Rapid acquisition of unknown DNA sequence adjacent to a known segment by multiplex restriction site PCR. *Biotechniques*, **25**: 415-419.
- West R.L. (1976a). Red to white fibre ratios as an index of double muscling in beef cattle. *Journal of Animal Science*,**38**: 1165-1175.
- West R.L.(1976b). Carcas traits and the organoleptic, chemical and histochemical characteristics of muscles from double muscled and normal cattle. *Dis. Abs. Int.*, B, **36**: 6051.
- Wrana J.L., Attisano L., Wieser R., Ventura F., Massague J. (1994). Mechanisms of activation of the TGF- $\beta$  receptor. *Nature*, **370**: 341-347.
- Wriedt C. (1929). Die Vererbung des Doppellendercharakters bei Rindern. *Z Indukt AbstamVererbungslebre*, **51**: 482-486.

## Anexo 1- Soluciones utilizadas.

**AHC** (500ml.), medio cultivo líquido para levaduras.

Glucosa	10 g.
<i>Yeast Nitrogen Base</i>	3,35 g.
<i>Casein Hydrol Acid</i>	5 g.
<i>Adenine Hemisulf (stock 0.5%)</i>	2 ml.
NaOH 2M	1ml.
H <sub>2</sub> O	Hasta 1 l.

\*Autoclavar

**Ampicilina**, solución *stock*.

50 mg./ml., disuelto en agua.

\*Guardar a -20 °C.

**Church** solución (1 l.)

Tampón fosfato 1M, pH=7,2	500 ml.
SDS (en polvo)	70 g.
BSA, fracción V (en polvo)	10 g.
EDTA 500mM., pH=8	2 ml.
H <sub>2</sub> O UP	hasta 1 l.

**5,7 M. CsCl**, tratado con DEPC.

Disolver CsCl en 0,1 M. EDTA, pH=8

Añadir 0,002 vol. DEPC, agitar 30' y autoclavar.

**Denhardt's 100x**, solución

<i>Bovine serum albumin</i>	2 g.
Ficoll TM 400	2 g.
Polyvinylpyrrolidona	2 g.
H <sub>2</sub> O	hasta 100 ml.

\*Conservar a -20°C. hasta 3 meses.

**Desnaturalización, solución**

NaCl (1.5 M.)	87,66 g.
NaOH (0.5 M.)	20 g.
H <sub>2</sub> O	hasta 1 l.

\*Conservar a temperatura ambiente, hasta 3 meses.

**EDTA 0,1M., pH=8. Solución (100 ml.)**

EDTA .2 H <sub>2</sub> O	3,72 g.
H <sub>2</sub> O	80 ml.
pH=8, con NaOH	
H <sub>2</sub> O	hasta 100 ml.

**ETM, solución (para 10 muestras)**

EDTA 0,5 M., pH=8	6,25 ml.
H <sub>2</sub> O	25 ml.
Tris HCl 1M., pH=8	312,5 µl.
β-mercaptoetanol	125 µl.

\*Se debe preparar en el momento.

**FELB**, (*Formamida EDTA Loading Buffer*), tampón de carga para muestras en geles de acrilamida.

Formamida	10 ml.
Xilencianol	10 mg.
Azul de bromofenol	10 mg.
EDTA 0,5 M.(pH=8)	200 µl.

**Hibridación, solución.**

Formamida desionizada	50 %
SSPE	5 x
Denhart's	5 x
SDS	0,1 %
ADN de espermatozoos de salmón desnaturalizado (10' a 100°C.)	100 µg./ml.

**IPTG** (200mg./ml.), solución

IPTG	1 g.
H <sub>2</sub> O UP	5 ml.

\*Conservar a -20° C.

**1 dilution buffer**, tampón para diluir y conservar fagos.

Tris-HCl 1M. pH= 7,5	10 ml.
MgSO <sub>4</sub> , 7 H <sub>2</sub> O	2g.
H <sub>2</sub> O	hasta 1l.

\* Autoclavar

**LB**, medio cultivo para bacterias.

Bactotryptona	10 g.
Extracto de levadura	5 g.
NaCl	10 g.
H <sub>2</sub> O	hasta 1 l.
pH= 7,0 con NaOH 5N.	

\* Autoclavar y guardar a 4°C.

**LB-agar**, medio de cultivo sólido para bacterias (1 l.)

Bactotryptona	10 g.
Extracto de levadura	5 g.
NaCl	10 g.
Agar	15 g.
H <sub>2</sub> O	hasta 1 l.
pH= 7,0 con NaOH 5N.	

\* Autoclavar.

**LB-agar-MgSO<sub>4</sub>** (10 mM.), medio de cultivo sólido para bacterias hospedadoras de fagos (1 l.).

Bactotryptona	10 g.
Extracto de levadura	5 g.
NaCl	10 g.
MgSO <sub>4</sub> . 7 H <sub>2</sub> O	2,4 g.
Agar	15 g.
H <sub>2</sub> O	hasta 1 l.
pH= 7,0 con NaOH 5N.	

\* Autoclavar.

**LB-MgSO<sub>4</sub> (10mM.), medio cultivo para bacterias hospedadoras de fagos.**

Bactotryptona	10 g.
Extracto de levadura	5 g.
NaCl	10 g.
MgSO <sub>4</sub> . 7 H <sub>2</sub> O	2,4 g.
H <sub>2</sub> O	hasta 1 l.
pH= 7,0 con NaOH 5N.	

\* Autoclavar y guardar a 4°C.

**LB-MgSO<sub>4</sub>-Maltosa**

Bactotryptona	10 g.
Extracto de levadura	5 g.
NaCl	10 g.
MgSO <sub>4</sub> . 7 H <sub>2</sub> O	2,4 g.
Maltosa	2 g.
H <sub>2</sub> O	hasta 1 l.
pH= 7,0 con NaOH 5N.	

\* Autoclavar y guardar a 4°C.

**LB-soft top agar (½ l.).**

Bactotryptona	5 g.
Extracto de levadura	2,5 g.
NaCl	5 g.
MgSO <sub>4</sub> . 7 H <sub>2</sub> O	1,2 g.
Agarosa	3,6 g..
H <sub>2</sub> O	hasta ½ l.
pH= 7,0 con NaOH 5N.	

\* Autoclavar y mantener a 50° C.

**LDS, solución (100ml.)**

EDTA 0,5 M., pH=8	20ml.
Tris 1M, pH=8	1ml.
H <sub>2</sub> O	hasta 100ml.
Añadir <i>Lithium Dodecyl Sulfate</i>	1g.

\*Se debe preparar en el momento.

**Lisis, solución.**

Tiocianato de Guanidina 4M.  
 NaAc 0,1 M., pH= 5,2  
 DEPC 0,1-0,2 %.  
 Autoclavar.  
 β-mercaptoetanol 2 %.

**Neutralización, solución.**

NaCl (1.5M.)	87,66 g.
Tris-HCl (0,5M.)	60,5 g.
EDTA (0,001M.)	0,37 g.
H <sub>2</sub> O	hasta 1 l.
pH=7,5 con HCl concentrado.	

\*Guardar a temperatura ambiente, hasta 3 meses.

**NZY, medio cultivo bacterias hospedadoras de fagos.**

NZ-amine	10 g.
NaCl	5 g.
Extracto de levadura	5 g.
MgSO <sub>4</sub> 7 H <sub>2</sub> O	2 g.
H <sub>2</sub> O	hasta 1 l.
pH=7,5 con NaOH	

\*Autoclavar y guardar a 4°C.

**NZY-agar, medio cultivo sólido bacterias hospedadoras de fagos.**

NZ-amine	10 g.
NaCl	5 g.
Extracto de levadura	5 g.
MgSO <sub>4</sub> 7 H <sub>2</sub> O	2 g.
Agar	15 g.
H <sub>2</sub> O	hasta 1 l.
pH=7,5 con NaOH	

\*Autoclavar.

**NZY-soft agar**

NZ-amine	10 g.
NaCl	5 g.
Extracto de levadura	5 g.



MgSO <sub>4</sub> 7 H <sub>2</sub> O	2 g.
Agarosa	7,2 g.
H <sub>2</sub> O	hasta 1 l.
pH=7,5 con NaOH	

\*Autoclavar y mantener a 50°C.

**SM buffer**, tampón para diluir y conservar fagos.

NaCl	5,8 g.
MgSO <sub>4</sub> 7 H <sub>2</sub> O	2g.
Tris-HCl 1M. , pH=7,5	50ml.
Solucion de gelatina 2%	5ml.
H <sub>2</sub> O	hasta 1l.

\* Autoclavar

**SSPE 20x**, solución

NaCl	175,3 g.
NaH <sub>2</sub> PO <sub>4</sub> , H <sub>2</sub> O	27,6 g.
EDTA 0,5 M.	40 ml.
H <sub>2</sub> O	hasta 1 l.
PH= 7,4 con NaOH	

**S.O.B.**, medio cultivo.

Bactotryptona	20 g.
Extracto de levadura	5 g.
NaCl	0,5 g.
KCl, 250mM.	10 ml.
H <sub>2</sub> O	hasta 1 l.
pH= 7,5 con NaOH 5N.	

\*Autoclavar

Antes de usar añadir: MgCl 2M. (estéril) 5 ml.

**S.O.C.**, medio cultivo

S.O.C. + glucosa 1M. (estéril)	20 ml.
--------------------------------	--------

**SPE**, solución

Sorbitol	1M.
NaH <sub>2</sub> PO <sub>4</sub>	1,6mM.

$\text{Na}_2\text{HPO}_4$	8,4mM.
EDTA	5mM., pH=8

**Spheroblasting**, solución (para 10 muestras)

SPE	1,2 ml.
<i>Zymolase</i>	1mg.
$\beta$ -mercaptoetanol	50 microl.

\*Se debe preparar en el momento.

**SSC 20x**, solución (1 l.)

Tri-sodio citrato	88,23 g.
NaCl	175,32 g.
$\text{H}_2\text{O}$	hasta 1 l.
pH entre 7 y 8	

**Tampón Fosfato**, solución (1 l.)

$\text{Na}_2\text{HPO}_4$ anhidro	142g
Disolver en 800 ml. $\text{H}_2\text{O}$	
Ajustar el pH a 7,2 con ácido fosfórico concentrado	
$\text{H}_2\text{O}$	hasta 1l.

**Tampón de hibridación**, utilizado en protocolos de *Primer Extension* (1 ml.).

PIPES	0,5 M.
EDTA	0,5 M.
NaCl	5 M.
Formamida desionizada	80 %
$\text{H}_2\text{O}$ DEPC.	hasta 1 ml.

**T<sub>10</sub>E<sub>1</sub>, pH=8**, solución (100 ml.)

Tris-HCl 1M pH= 8	10 ml.
EDTA 100mM. pH= 8	10 ml.
$\text{H}_2\text{O}$	hasta 1 l.

\*Autoclavar y guardar a temperatura ambiente.

**TBE 10x**, solución (1 l.)

Tris Base	109 g.
-----------	--------

Ácido bórico	55 g.
EDTA 0,5M., pH= 8,0	40 ml.
H <sub>2</sub> O	hasta 1 l.

**Tris HCl 1M pH= 8, solución (1 l.)**

Tris-base	121,1 g.
H <sub>2</sub> O	800 ml.
pH=8 con HCl 32%	> 60 ml.
H <sub>2</sub> O	hasta 1 l.

\*Autoclavar y mantener a temperatura ambiente.

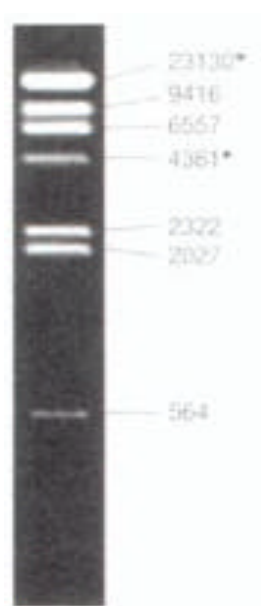
**X-Gal (20mg./ml.), solución en Dimetyl-formamida**

X-Gal	100 mg.
Dimetyl-formamida	5 ml.

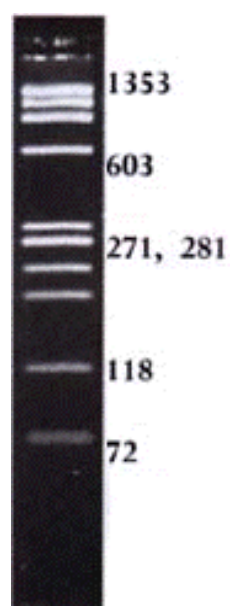
\*Conservar a -20° C.

## Anexo 2- Marcadores de tamaño de ADN.

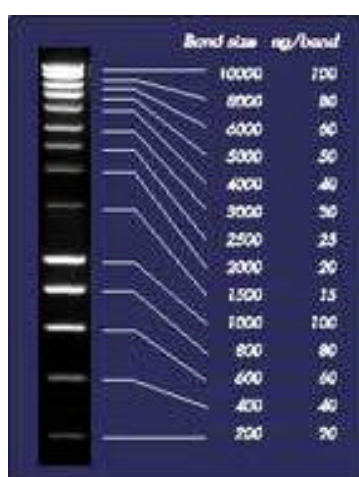
-  $\phi$  / *Hind III* (Pharmacia)



-  $\phi$ X174 RF DNA *Hae III* digest (Pharmacia)



- Smart Ladder (Eurogentec)



- **Anexo 3. Secuencias de la miostatina bovina.**

En negro secuencias no transcritas. En azul las secuencias de los exones no traducidas a proteínas. En rojo la secuencia traducida, y su traducción en proteínas. Subrayado el péptido señal y el lugar de procesamiento proteolítico, además de los residuos Cisteína conservados en la familia TGF- $\beta$ . Y los lugares de N-glicosilación de la proteína. En verde las regiones CAAT-box y la TATA-box, los límites de los intrones GT-AG, y lugares de poliadenilación. En azul claro las zonas dianas del procesamiento o maduración de los intrones.

**a) Región del promotor y exón1.**

...4857.GCGAGATTTCATTGTGGAGCAAGAG**CAAT-Box**CCAGATCCCGACGACACTTGTCTCATCAAAGTT**TATA-Box**GAATATAAGCCACTTGAATACAGTATA**Inicio2**AAAA**McPherron**GATTCACTGGTGTGGCA**Iniciol**AGTTGTCTCTCAGACTGGGCAGGCATTAACTTTG

...5007.GCTTGGCGTTACTCAAAGCAAAGAAAAGTAAAGGAAAAAGTAAGAACAAGGGAAAAGATTGTATTGA.TTTTAAAC**Péptido señal**ATGCAAAACTGCAAACTCTGTGTTTATATTACCTATTTATGCTGATTGTTGCTGGCCAGTGGATCTG

**N-Glyc** N E N S E Q K E N V E K E G L C N A C L W R E N T T S S R L E A I K I Q I L S K L R L E T A P N I S **N-Glyc**

...5156.AATGAGAACAGCGAGCAGAAGGAAATGTGGAAAAAGAGGGGCTGTGTAATGCATGTTTGTGGAGGAAAACACTACATCCTCAAGACTAGAAGCCATAAAATCCAAATCCTCAGTAAACTTCGCCTGGAACAGCTCCTAACATCAGC

K D A I R Q L L P K A P P L L E L I D Q F D V Q R D A S S D G S L E D D D Y H A R T E T V I T M P T

...5306.AAAGATGCTATCAGACAACTTTTGCCCAAGGCTCCTCCACTCCTGGAAGTGAATGATCAGTTTCGATGTCCAGAGAGATGCCAGCAGTGACGGCTCCTTGAAGACGATGACTACCACGCCAGGACGAAACGGTCATTACCATGCCACG

...5456.GAGTGTGAGTAGTCTGCTGGTGCAAGCAACGACTCTGCTGACTGCTGTTCTAGTGTTCATGAAAAACCGATCTATTTTCAGGCTCTTTTAAACAGCTGCTGGCTTGTATGTAAGGAGGAGGGGAAAGAGCTTTTTT...CAAGAT

**b) Exón 2.**

...7086.TAAAAGATTCTGAAAAGCTGTAATAACTGTTATACTTGATAT...TTTGCTGTATGAATGAAATGCTACATATTTTCCATTTTAAAAGACTAAATATGCACACATTATTCCAATTAAAAATGTTTCATAGA**TTGAT**ATGGAGGTGTTT

S D L L T Q V E G K P K C C F F K F S S K I Q Y N K L V K A Q L W

...7233.GTTTGTTTTTCATAAAATGATCTTAGTAAGTCTTTTCTTATTTCATTTATAGCTGATCTTCTAACGCAAGTGAAGGAAAACCCAAATGTTGCTTCTTTAAATTTAGCTCTAAGATACAATAACAATAAGTAAAGGCCCACTGTG

I Y L R P V K T P A T V F V Q I L R L I K P M K D G T R Y T G I R S L K L D M N P G T G I W Q S I D

...7383.GATATATCTGAGGCCTGTCAAGACTCCTGCGACAGTGTGTTGTGCAATCCTGAGACTCATCAAACCATGAAAGACGGTACAAGGTATACTGGAATCCGATCTCTGAACTTGACATGAACCCAGGCACTGGTATTGGCAGAGCATTGA

V K T V L Q N W L K Q P E S N L G I E I K A L D E N G H D L A V T F P E P G E D G L

...7533.TGTGAAGACAGTGTGTCAGAACTGGCTCAAACAACCTGAATCCAACCTAGGCATTGAAATCAAAGCTTTAGATGAGAATGGCCATGATCTTGCTGTAACTTCCAGAACCCAGGAGAAGATGGACTG**GT**AAGTGATTACTGAAAATAACA

### c) Exón 3.

Lugar de procesado proteolítico

...9647..AGTGTTTTGGGATCTATTA**TTAAC**CTTCTTTTCCTTTCCATAC**ACTCCTTTTTTAGAAGTCAAGGTAACAGACACACCAAAAAGATCTAGGAGAGATTTTGGGCTTGATTGTGATGAACACTCCACAGAATCTCGATGCTGCTGTTAC**

P L T V D F E A F G W D W I I A P K R Y K A N Y **C S G E C E F V F L Q K Y P H T H L V H Q A N P R G**

...9797..**CCTCTAACTGTGGATTTTGAAGCTTTTGGATGGGATTGGATTATTGCACCTAAAAGATATAAGGCCAATTACTGCTCTGGAGAATGTGAATTTGTATTTTGC AAAAGTATCCTCATACCCATCTTGTGCACCAAGCAAACCCAGAGGT**

S A G P **C C** T P T K M S P I N M L Y F N G E G Q I I Y G K I P A M V V D R **C G C S \***

...9947..**TCAGCCGGCCCTGCTGTACTCCTACAAAGATGTCTCCAATTAAATATGCTATATTTTAAATGGCGAAGGACAAATAATATACGGAAGATTCCAGCCATGGTAGTAGATCGCTGTGGGTGTTTCATGAGGTCTATATTTGGTTCATAGCTTC**

...10097..CTCAAACATGGAAGGTCTTCCCTCAACAATTTTGAACTGTGAAATTATGTACCACAGGCTATAAGCCTAGAGTAGTGTACAGTCACTTAAGCACAAGCTACAGTATATGAGCTAAAAGAGAGAATATATGCAATGGTTGGCATTAA

...10247..CCATCCAAACAAATCGTATAATAAAAAGTTTATGATTTCAGAGTTTTTGAACATA...GGAGATCAAATTCATTTATGTTGAAATAT.....ATTACAACACATGCAGGTGAATGAAAGCAATTCTCCTTGTCTTCTGGTGAA

...10384..TTAAAGGAGTATGCTTTAAA.....ATCTATTTCTCTACAGTTTCACTTAATATTTACAGAAAAATCTATATGTAGTATTGGTAAAATGCAGTATTGTTATATACCATTATTTGAAACATCCTTAAACACTTGAATTTATATTGTATG

...10527..ATAGCATACTTGGTAAGATGAGATTCCACAAAGTAGGGATGGCACACCATACGCAGGTTACCATTCTTATACTGATTGATACAGTACATTAAACAGTTTTTGCCAATGGTGCATAACAATAGGCTGAATGGCTGATGTTATCAGGTTTAT

LP4

...10677..CAAG**CAAAAAA**CATTTCAGGAAAGTAATAAGTTTCTCCTTTCTTCAGGTGCATTTTCACACTCCTCCCTA..TGGGCAATGGATGTTCTATAAAGAAAGAAAATC..ATTTTCTAGAGGTCTACATTCAAT..TCTGTAGCATACTTGA

...10822..GAAGCTGC...ATTGAAAAGGCAGTCAAAAAGTATTCA..TTTTGGTCAAAAATTTCAAAAATTATAGCCTGCCTTTGCAATACTGCAGCTTTTAGGATGAAATAATGGAAATGACTGATTCTATCAATATTGTATAAAAAGATTTTGAAA

LP3

...10966..TAGTTGCATTTATATAATATGTATACAATATTGTTTTGTAA**ATAAA**TGTCTCCTTTTTTATT**TA**CTTTGGTATATTTTACAGTAAGGACATTTCAAATTAAGTATTAAGGCACAAAGACATGTCTAGTATAGGACATAAAA..GCAAAAGCT

...11115..TATATTTTGGAGCAAATTAGTTGATTAAATAGTGGTCTTAAAACTCCATATGCTAATGGTTAGATGGTTATATTACAATCATTTTATAT..TTTTTACATTATTAGCATTCACTTATGGATTTCATGATGGCTGTATAATGTGAATGTGAA

...11264..ATTTCAATGGTTTACTGTCAATTGTATTCAAATCTCAACGTTCCATTATTTTAATACTTATAAATATTAAGCATACCAAAATGATTTAACTCTATTATCTGAAATCAGAATAATAAACTGATGATATCTTACAAATTGTTA..ATTTTATT

LP2

...11412..TTATAATTTGATAATGAATATATTTCTGCATATATTTACTACTATTTTGTAATCAGGATTTTGTTAATCA**ATAAA**TTGTACTTATGATTAAGT**GAAA**TTATTTCTTACAT..CCAATGTGTAGAAACAATTTAAGTTAT**ATTAAAG**GTG

LP1

...11561..TTTCACCTTTTTTGAA**AGACAA**CAGTTTCATGTTATAATGATTAACTCTAGATTTCT..GGTTTCCACTTTATTATAAAAGTTT..AATGACTGAGCACAAAAGTTTGGTTTAGAAATTTTAGGTCTGCTACTCTAGTTTCTCATGAGTG

#### **Anexo 4- Publicaciones más relevantes.**

---

- L. Grobet, **L.J. Royo**, D. Poncelet, D. Pirottin, B. Brouwers, J. Riquet, A. Schoberlein, S. Dunner, F. Menissier, J. Massabanda, R. Fries, R. Hanset, M. Georges (1997). A deletion in the bovine myostatin gene caused the double-muscled phenotype in cattle. *Nature Genetics*, **17**: 71-74.
  
- L. Grobet, D. Poncelet, **L.J. Royo**, B. Brouwers, D. Pirottin, C. Michaux, F. Menissier, M. Zanotti, S. Dunner, M. Georges (1998). Molecular definition of an allelic series of mutations disrupting the myostatin function and causing double muscled in cattle. *Mammalian Genome*, **9**: 210-213.
  
- D. Pirottin, D. Poncelet, L. Grobet, **L.J. Royo**, B. Brouwers, J. Masabanda, H. Takeda, R. Fries, K. Sugimoto, J.E. Womack, S. Dunner, & M. Georges (1999). High resolution human-bovine comparative mapping based on a closed YAC contig spanning the bovine mh locus. *Mammalian Genome*, **10**, 289-293.



# A deletion in the bovine myostatin gene causes the double-muscled phenotype in cattle

Luc Grobet<sup>1</sup>, Luis José Royo Martín<sup>2</sup>, Dominique Poncelet<sup>1</sup>, Dimitri Pirottin<sup>1</sup>, Benoit Brouwers<sup>1</sup>, Juliette Riquet<sup>1</sup>, Andreina Schoeberlein<sup>1</sup>, Susana Dunner<sup>2</sup>, François Ménissier<sup>3</sup>, Julio Massabanda<sup>4</sup>, Ruedi Fries<sup>4</sup>, Roger Hanset<sup>1</sup> & Michel Georges<sup>1</sup>

An exceptional muscle development commonly referred to as 'double-muscled' (Fig. 1) has been seen in several cattle breeds and has attracted considerable attention from beef producers. Double-muscled animals are characterized by an increase in muscle mass of about 20%, due to general skeletal-muscle hyperplasia—that is, an increase in the number of muscle fibers rather than in their individual diameter<sup>1</sup>. Although the hereditary nature of the double-muscled condition was recognized early on, the precise mode of inheritance has remained controversial; monogenic (dominant and recessive), oligogenic and polygenic models have been proposed<sup>2</sup>. In the Belgian Blue cattle breed (BBCB)<sup>4</sup>, segregation analysis performed both in experimental crosses<sup>3</sup> and in the outbred population suggested an autosomal recessive inheritance. This was confirmed when the muscular hypertrophy (*mh*) locus was mapped 3.1 cM from microsatellite TGLA44 on the centromeric end of bovine chromosome 2 (ref. 5). We used a positional candidate approach to demonstrate that a mutation in bovine *MSTN*, which encodes myostatin, a member of the TGF $\beta$  superfamily, is responsible for the double-muscled phenotype. We report an 11-bp deletion in the coding sequence for the bioactive carboxy-terminal domain of the protein causing the muscular hypertrophy observed in Belgian Blue cattle.

Previous work has shown that the pro- $\alpha$ (III) collagen gene (*Col3A1*) is located in the same chromosomal region as the *mh* locus. *Col3A1* has been mapped to BTA2q12–22 by *in situ* hybridization<sup>6</sup>, while a *Col3A1* RFLP marker was shown to be closely linked to TGLA44 ( $\theta=2\%$ )<sup>7</sup>. This identifies the region flanking *Col3A1* on the human map—HSA2q31–33—as the likely orthologous human chromosome segment, which is compatible with Zoo-FISH data<sup>6</sup> and mapping of type-I markers on somatic cell hybrids<sup>8</sup>.

To refine the correspondence between the HSA2q31–33 and BTA2q12–22 maps, we developed comparative anchored tagged sequences (CATS)—primer pairs that would amplify an STS from the orthologous gene in different species<sup>9</sup>—for a series of genes flanking *Col3A1* on the human map and for which sequence information was available in more than one mammal. In addition to *Col3A1*, CATS were obtained for *Col5A2*, *INPP1*, *TFPI*, *TTN*, *CHN*, *GAD1*, *CTLA4* and *CD28*. We used these CATS, as well as all microsatellite markers available for proximal BTA2—TGLA44, BM81124, BM3627, ILSTS026, INRA40 and TGLA431 (ref. 10)—to screen a six-genome equivalent bovine

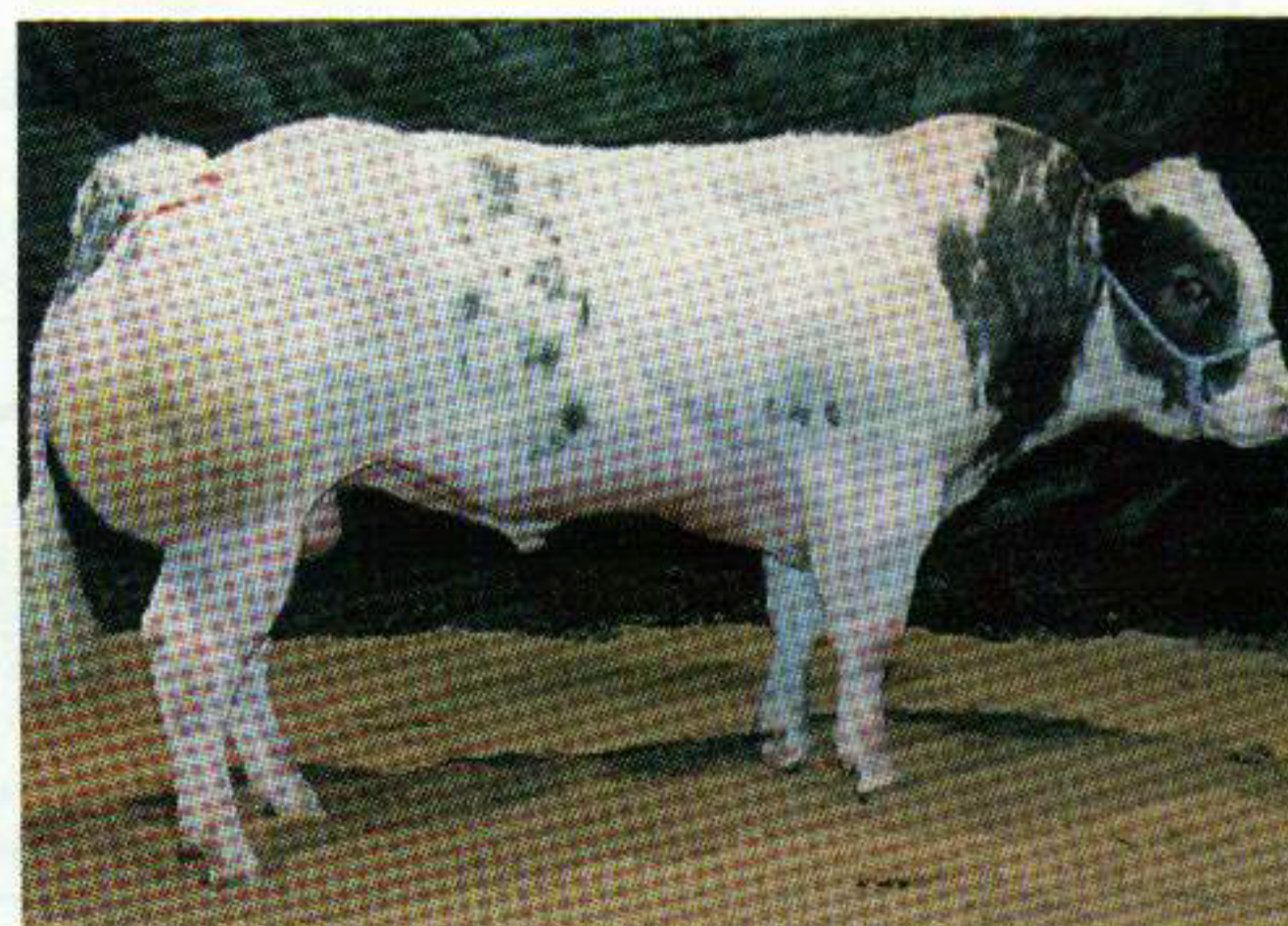


Fig. 1 Double-muscled Belgian Blue animal homozygous for the *nt821del(11)* deletion in the myostatin gene.

YAC library (A. Schoeberlein, manuscript in preparation). We explored potential overlap between YACs obtained with this panel of STSs on the basis of common STS content, as well as cross-hybridization between individual SINE-PCR product. This analysis revealed three independent YAC contigs in the region of interest (Fig. 2): i) contig A, containing microsatellites TGLA44, BM81124 and *INPP1*; ii) contig B, containing *Col3A1* and *Col5A2*; and iii) contig C, containing microsatellite markers BM3627, ILSTS026, INRA40 and *TFPI*. We isolated new polymorphic microsatellite markers from contig B (BULGE20), as well as contig A (BULGE23), in order to increase the local information content. We genotyped a previously described experimental backcross population<sup>5,15</sup>, and used all available genotypes to construct a linkage map with the ILINK program. We obtained the following most likely order and sex-averaged recombination rates between adja-

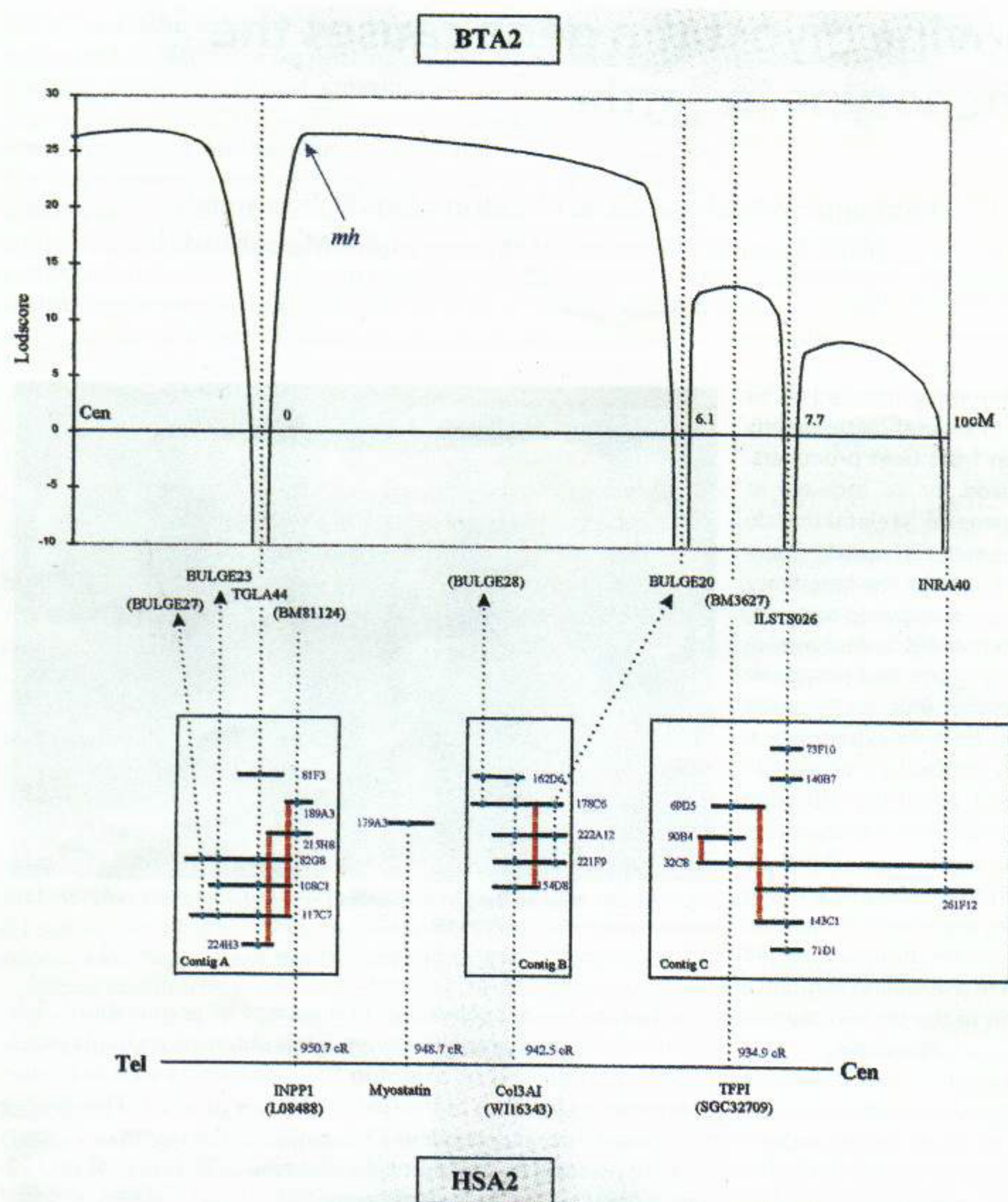
Table 1 • Distribution of *nt821(del11)* mutation in different cattle breeds

Breed	Phenotype	Genotype		
		+/+	+/ <i>nt821(del11)</i>	<i>nt821(del11)/nt821(del11)</i>
Belgian Blue	DM	—	—	104
Asturiana	DM	—	—	52
Maine-Anjou	DM	42	—	—
Holstein-Friesian	Normal	100	—	—
Jersey	Normal	96	—	—

<sup>1</sup>Department of Genetics, Faculty of Veterinary Medicine, University of Liège (B43), 20 Bd de Colonster, 4000 Liège, Belgium. <sup>2</sup>Laboratorio de Genética, Dpto. de Producción Animal, Universidad Complutense, 28040 Madrid, Spain. <sup>3</sup>Station de Génétique Quantitative et Appliquée, I.N.R.A., 78352 Jouy-en-Josas, France.

<sup>4</sup>Technische Universität München, D-85350 Freising-Weihenstephan, Germany. Correspondence should be addressed to M.G. e-mail: michel@stat.fmv.ulg.ac.be





**Fig. 2** Summary of genetic, physical and comparative mapping data around the bovine *mh* locus. A multipoint lod-score curve obtained for the *mh* locus with respect to the microsatellite marker map is shown. Markers that were not informative in our pedigree are shown between brackets; their map position is inferred from published mapping data. Markers and the YACs from which they were isolated are connected by arrows. The RH map of the relevant section of human HSA2 is shown, with the relative position in cR of the ESTs used in this work. Stippled lines connect microsatellite and type I markers with their respective positive YACs. YACs showing cross-hybridizing SINE-PCR products are connected by the red boxes.

cent markers: [TGLA44-(0%)-BULGE23]-(6.1%)-BULGE20-(1.6%)-ILSTS026-(2.3%)-INRA40-7.1%-TGLA431. The position of BULGE20 between TGLA44 and ILSTS026 confirmed the order of the three contigs, which had been predicted from available human mapping data (Fig. 2).

The *mh* locus was placed on this new marker map by multipoint linkage analysis, yielding a lod-score curve (Fig. 2) and positioning the *mh* locus in the TGLA44-BULGE20 interval, with an associated maximum lod score of 26.4. Given the relative position of these microsatellite markers with respect to *INPP1* and *Col3A1*, as deduced from the integration of the human and bovine map, these results indicated that the *mh* gene was likely to be located in a chromosome segment bounded by *INPP1* and *Col3A1*.

Recently, McPherron *et al.*<sup>11</sup> demonstrated that mice homozygous for a knockout deletion of *Mstn* (formerly called *Gdf8*), encoding myostatin, were characterized by a general increase in skeletal-muscle mass, similar to the bovine double-muscling phenotype. Using the published murine *Mstn* cDNA sequence, we iden-

untranslated sequence, respectively, in two overlapping DNA fragments of 753 (primers GDF8.27-GDF8.12) and 724 bp (primers GDF8.11-GDF8.21) long. We successfully amplified the expected DNA products from cDNA generated from skeletal muscle of both a wild-type (homozygous +/+) and a double-muscling (homozygous *mh/mh*) animal, and sequenced them on both strands.

The nucleotide sequence of the wild-type allele shows 89.1% identity with the mouse *Mstn* coding sequence over a 1,128-bp overlap. The predicted protein shows 92.5% identity in a 375-amino-acid overlap with mouse myostatin. As expected for a member of the *TGFβ* superfamily, the bovine myostatin gene is characterized by a proteolytic processing site thought to mediate cleavage of the bioactive carboxy-terminal domain from the longer N-terminal fragment, and by nine cysteine residues separated by a characteristic spacing and believed to be involved in intra- and inter-molecular disulphide bridges<sup>14</sup>.

The nucleotide sequence obtained from the *mh* allele was identical to the wild-type allele, except for an 11-bp deletion in nucleotides

tified a tentative human consensus (THC) cluster in the Unigene database representing three cDNA clones (221299, 300367, 308202) and six EST sequences (H92027, H92028, N80248, N95327, W07375, W24782). The corresponding THC covered 1,296 bp of human cDNA, showing a homology of 78.1% with murine myostatin when averaged over the entire sequence and 91.1% when only the translated parts of the human and murine genes (566 bp) were considered. This THC therefore clearly corresponded to the human orthologue of the murine myostatin gene. We designed primers to amplify a 272-bp fragment from the second exon of the human myostatin gene (*MSTN*) and to genotype the Genebridge-4 radiation hybrid panel<sup>12</sup>. We placed *MSTN* on the Whitehead/MIT framework radiation hybrid map, at position 948.7 cR of the HSA2 map with an associated lod score greater than 3. Closer examination of the *MSTN* segregation vector and comparison with the vectors from all markers located in that region showed it to be identical to EST SGC38239, placed on the Whitehead/MIT radiation hybrid map<sup>13</sup> at position 946.8 cR of HSA2. This places *MSTN* in the interval between *Col3A1* (EST WI16343-942.5 cR) and *INPP1* (EST L08488-950.2-951.2 cR; Fig. 2), making bovine *MSTN* a very strong positional candidate for the *mh* gene.

To test whether myostatin is associated with the double-muscling phenotype in cattle, we designed primer pairs based on the available mouse and human sequences, aiming for RT-PCR amplification of the entire coding sequence from bovine cDNA. We identified two primer pairs that jointly amplified the entire coding sequence as well as 76 bp and 83 bp of 5' and 3'







BBCB<sup>5</sup> and Asturiana<sup>15</sup> was reported previously. Linkage analysis, performed in Maine-Anjou pedigrees segregating for the double-muscling phenotype, yielded a maximum lod score of 5.2 at 6.4% recombination from TGLA44, clearly involving the same chromosomal region and probably the *mh* gene in this breed as well. All 196 dairy animals were homozygous for the wild-type allele. In BBCB and Asturiana, all double-muscling animals were homozygous for the *nt821(del11)* deletion, demonstrating allelic homogeneity in these two breeds. In contrast, all double-muscling Maine-Anjou were homozygous for the wild-type allele and did not carry the *nt821(del11)* deletion. The most likely explanation for this finding would be allelic heterogeneity of the bovine myostatin gene. If confirmed, this would force re-evaluation of the popular hypothesis that a unique mutation was disseminated across Europe with the once popular Shorthorn breed at the beginning of the 19th century<sup>2</sup>. Allelic heterogeneity in the bovine myostatin gene might also account for the conflicting genetic models proposed for the inheritance of the trait in different breeds. As members of the TGF $\beta$  superfamily are known to act as dimers, dominant negative mutations could account for the dominant inheritance observed for the double-muscling condition in some breeds.

The identification of *MSTN* as the gene causing the double-muscling phenotype will allow for the development of diagnostic tests that will facilitate the selection for or against this trait in cattle. It also paves the way towards manipulation of muscle development in livestock by means of various approaches, including immunomodulation and transgenesis. It would be interesting to determine whether inactivation of *MSTN* after birth, using a *Cre-loxP* system or related approach, might still lead to increased muscle development. Moreover, the identification of the myostatin gene as a key regulator of muscle development will permit study of upstream and downstream factors (such as the myostatin receptor) that might lead to the identification of other genes underlying genetic variation for muscle development in livestock.

## Methods

**Genetic mapping.** Genetic mapping was performed in a previously described (Holstein-Friesian x Belgian Blue) x Belgian Blue experimental backcross population with 108 informative offspring<sup>5</sup>. The pedigree material available in Maine-Anjou was composed of five paternal half-sib pedigrees comprising a total of 42 double-muscling offspring from five conventional sires. Microsatellite genotyping was performed according to standard procedures<sup>16</sup>, using the primer sequences reported in Table 2.

Linkage analyses were performed with the MLINK, ILINK and LINKMAP programs of LINKAGE (version 5.1; ref. 17) and FASTLINK (2.3P version, June 1995; ref. 18) packages. Linkage analysis of the *mh* locus was performed under a simple recessive model assuming full penetrance for *mh/mh* individuals and zero penetrance for the two other genotypes.

**Screening of the YAC library.** The yeast artificial chromosome library was screened by PCR a three-dimensional pooling scheme as previously described<sup>19</sup>. The primer pairs corresponding to the CATS used to screen the library are reported in Table 2. Cross-hybridization between SINE-PCR products of individual YACs was performed as described (using primers reported in ref. 21)<sup>20</sup>. Microsatellites were isolated from YACs as described<sup>22</sup>.

**Mapping of the human myostatin gene.** DNA from the Genebridge-4 panel<sup>12</sup> was purchased from Research Genetics (Huntsville, AL) and genotyped by PCR according to standard procedures and the following primer pair corresponding to the human myostatin gene sequence: 5'-GGCC-CAACTATGGATATATTTG-3' and 5'-GGTCCTGGGAAGGTTAC-AGCA-3'. Mapping was performed with the WWW server of the Whitehead Institute/MIT Center for Genome Research and their *RH-mapper* program (Slonim, D., Stein, L., Kruglyak, L. & Lander, E., unpublished data) to position the markers with respect to the framework map. Segregation vectors of the query markers were compared with the vectors from all markers in the region of interest in the complete Data Release 11.9 (May 1997) to obtain a more precise position.

**RT-PCR.** Total RNA was extracted from skeletal muscle (triceps brachialis) as described<sup>23</sup>. RT-PCR was performed with the Gene-Amp RNA PCR Kit (Perkin Elmer) and the primers reported in Table 2. The PCR products were purified with QiaQuick PCR Purification kit (Qiagen) and sequenced with Dye terminator Cycle Sequencing Ready Reaction (Perkin Elmer) and an ABI373 automatic sequencer, using the primers reported in Table 2.

## Acknowledgements

This work was supported by a grant from the Belgian Ministère de l'Agriculture et des Classes Moyennes (D1/2-5744A) and a grant from the European Union (B104-CT95-0073). We are grateful to J. Weissenbach and C. Fizames for communicating the segregation vector of the L08488 EST in the GeneBridge-4 panel, and to D. Sereteyn, I. Dufrasne, D. Desmecht and F. Rolin for helping us to obtain the muscle biopsies. We acknowledge Holland Genetics and Livestock Improvement Corporation for providing us with the samples from Holstein-Friesian and Jersey animals. We thank P. Leroy for his continuous support and for funding the construction of the YAC library.

Received 2 July 1997; accepted 14 July 1997.

- Hanset, R. The major gene of muscular hypertrophy in the Belgian Blue cattle breed. in *Breeding for Disease Resistance in Farm Animals* (eds Owen, J.B. & Axford, R.F.E.) 467-478 (C.A.B. International, Oxford, 1991).
- Ménissier, F. Present state of knowledge about the genetic determination of muscular hypertrophy or the double muscling trait in cattle. in *Current Topics in Veterinary Medicine and Animal Science*, vol 16: *Muscle Hypertrophy of Genetic Origin and Its Use to Improve Beef Production* (eds King, J.W.B. & Ménissier, F.) 387-428 (Martinus Nijhoff, Norwell, Massachusetts, 1982).
- Hanset, R. & Michaux, C. On the genetic determinism of muscular hypertrophy in the Belgian White and Blue cattle breed. I. Experimental data. *Genet. Sel. Evol.* **17**, 359-368 (1985).
- Hanset, R. & Michaux, C. On the genetic determinism of muscular hypertrophy in the Belgian White and Blue cattle breed: II. Population data. *Genet. Sel. Evol.* **17**, 369-386 (1985).
- Charlier, C. et al. The *mh* gene causing double-muscling in cattle maps to bovine chromosome 2. *Mamm. Genome* **6**, 788-792 (1995).
- Solinas-Toldo, S., Lengauer, C. & Fries, R. Comparative genome map of man and cattle. *Genomics* **27**, 489-496 (1995).
- Fisher, S.R., Beever, J.E. & Lewin, H.A. Genetic mapping of COL3A1 to bovine chromosome 2. *Mamm. Genome* **8**, 76-77 (1996).
- O'Brien, S.J. et al. Anchored reference loci for comparative genome mapping in mammals. *Nature Genet.* **3**, 103-112 (1993).
- Lyons, A.L. et al. Comparative anchor tagged sequences (CATS) for integrative mapping of mammalian genomes. *Nature Genet.* **15**, 47-56 (1996).
- Kappes, S.M. et al. A second-generation linkage map of the bovine genome. *Genome Res.* **7**, 235-249 (1997).
- McPherron, A.C., Lawler, A.M. & Lee, S.-J. Regulation of skeletal muscle mass in mice by a new TGF $\beta$  superfamily member. *Nature* **387**, 83-90 (1997).
- Walter, M.A., Spillet, D.J., Thomas, P., Weissenbach, J. & Goodfellow, P.N. A method for constructing radiation hybrid maps of whole genomes. *Nature Genet.* **7**, 22-28 (1994).
- Hudson, T.G. et al. An STS-based map of the human genome. *Science* **270**, 1945-1954 (1995), with supplementary data from the Whitehead Institute/MIT Center for Genome Research, Human Genetic Mapping Project, data release 11.9 (May 1997).
- McPherron, A.C. & Lee, S.-J. The transforming growth factor  $\beta$  superfamily. in *Growth Factors and Cytokines in Health and Disease*, vol 1B (eds LeRoith, D. & Bondy, C.) 357-393 (Jai Publishing, Greenwich, Connecticut, 1996).
- Dunner, S. et al. Towards interbreed IBD fine mapping of the *mh* locus: double-muscling in the Asturiana de los Valles breed involves the same locus as in the Belgian Blue cattle breed. *Mamm. Genome* **8**, 430-435 (1997).
- Georges, M. et al. Mapping quantitative trait loci controlling milk production by exploiting progeny testing. *Genetics* **139**, 907-920 (1995).
- Lathrop, M. & Lalouel, J.M. Easy calculations of lodscores and genetic risk on small computers. *Am. J. Hum. Genet.* **36**, 460-465 (1984).
- Cottingham, R.W., Jr., Idury, R.M. & Schäffer, A.A. Faster sequential genetic linkage computations. *Am. J. Hum. Genet.* **53**, 252-263 (1993).
- Libert, F., Lefort, A., Okimoto, R., Womack, J. & Georges, M. Construction of a bovine genomic library of large yeast artificial chromosome clones. *Genomics* **18**, 270-276 (1993).
- Hunter, K.W. et al. Toward the construction of integrated physical and genetic maps of the mouse genome using interspersed repetitive sequence PCR (IRS-PCR) genomics. *Genome Res.* **6**, 290-299 (1996).
- Lenstra, J.A., van Bostel, J.A.F., Zwaagstra, K.A. & Schwerin, M. Short interspersed nuclear element (SINE) sequences of the *Bovidae*. *Anim. Genet.* **24**, 33-39 (1993).
- Cornélis, F. et al. Identification of a CA repeat at the TCRA locus using yeast artificial chromosomes: a general method for generating highly polymorphic markers at chosen loci. *Genomics* **13**, 820-825 (1992).
- Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. & Rutter, W.J. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* **18**, 5294-5299 (1979).



# Molecular definition of an allelic series of mutations disrupting the myostatin function and causing double-muscling in cattle

Luc Grobet,<sup>1,\*</sup> Dominique Poncelet,<sup>1,\*</sup> Luis José Royo,<sup>2</sup> Benoit Brouwers,<sup>1</sup> Dimitri Pirottin,<sup>1</sup> Charles Michaux,<sup>1</sup> François Méniissier,<sup>3</sup> Marta Zanotti,<sup>4</sup> Susana Dunner,<sup>2</sup> Michel Georges<sup>1</sup>

<sup>1</sup>Department of Genetics, Faculty of Veterinary Medicine, University of Liège (B43), 20 Bd de Colonster, 4000-Liège, Belgium

<sup>2</sup>Laboratorio de Genética, Dpto. de Producción Animal, Universidad Complutense, 28040 Madrid, Spain

<sup>3</sup>Station de Génétique Quantitative et Appliquée. I.N.R.A., 78352 Jouy-en-Josas, France

<sup>4</sup>Istituto di Zootecnia, facoltà di Medicina Veterinaria, Università degli Studi di Milano, 20133 Milano, Italy

Received: 20 September 1997 / Accepted: 11 October 1997

**Abstract.** We have determined the entire myostatin coding sequence for 32 double-muscled cattle sampled from ten European cattle breeds. Seven DNA sequence polymorphisms were identified, of which five would be predicted to disrupt the function of the protein, one is a conservative amino acid substitution, and one a silent DNA sequence variant. Four additional DNA sequence polymorphisms were identified in myostatin intronic sequences. In all but two breeds, all double-muscled animals were either homozygous or compound heterozygotes for one of the five loss-of-function mutations. The absence of obvious loss-of-function mutations in the coding sequence of the two remaining breeds points either towards additional mutations in unexplored segments of the gene, or towards locus heterogeneity of double-muscling.

## Introduction

The occurrence of individuals characterized by an inherited, exceptional muscularity—commonly referred to as double-muscling—has been reported in several cattle breeds (Culley 1807). The calving difficulties associated with double-muscling have led several breeding organizations to treat this condition as a genetic defect that needs to be eliminated. In specific economic contexts, however, the gains in feed conversion ratio, dressing out percentage, and meat quality (increased lean and tenderness) have outweighed the costs of dystocia, leading either to a systematic selection for double-muscled animals or their use in cross-breeding (Hanset 1991).

While the hereditary nature of this condition was recognized early on, the precise mode of inheritance has remained controversial. Contradictory inheritance models, including autosomal monogenic (dominant and recessive), oligogenic and polygenic, have been described (Méniissier 1982). It is at present unclear whether this reflects genuine genetic heterogeneity for this trait in different populations or whether it results from the difficulty to correctly classify animals for a trait that is in many respects a continuously distributed phenotype rather than a discrete affection status.

In the Belgian Blue Cattle Breed (BBB), segregation analysis performed both in experimental crosses and in the outbred population suggested an autosomal recessive inheritance (Hanset and Michaux, 1985a, & 1985b). This hypothesis was confirmed when the postulated *mh* locus (muscular hypertrophy) was unambiguously mapped in this breed to the centromeric tip of bovine Chromosome (Chr) 2 under a recessive model (Charlier et al., 1995).

The same chromosomal region was subsequently shown to account for the recessive inheritance of the double-muscling condition in at least two other breeds: Asturiana (Dunner et al. 1997) and Maine-Anjou (Grobet et al. unpublished). Its role in the determinism of double-muscling in other populations, however, remained to be demonstrated.

Recently, it was shown that an 11-bp deletion (*nt821(del11)*) in the myostatin gene, resulting in a truncation of the bioactive carboxyterminal domain of the protein, was causing double-muscling in BBB (Grobet et al. 1997). The same *nt821(del11)* mutation was shown to be responsible for double-muscling in Asturiana as well. In the Maine-Anjou breed, however, in which the involvement of the same *mh* locus was demonstrated by linkage analysis, the *nt821(del11)* deletion proved to be absent among double-muscled animals (Grobet et al. 1997). This observation, therefore, pointed towards allelic heterogeneity for the myostatin gene. This presumption was confirmed with the recent identification of a distinct cysteine to tyrosine substitution (C313Y) in double-muscled Piedmontese animals (Kambadur et al. 1997).

To further clarify the issue of locus and allelic heterogeneity of double-muscled cattle, we have determined the entire coding sequence of the myostatin gene in double-muscled individuals from ten European cattle breeds. In so doing, we have identified a series of mutations that are predicted to disrupt the myostatin function, demonstrating and characterizing the allelic heterogeneity of the double-muscled condition in cattle.

## Materials and methods

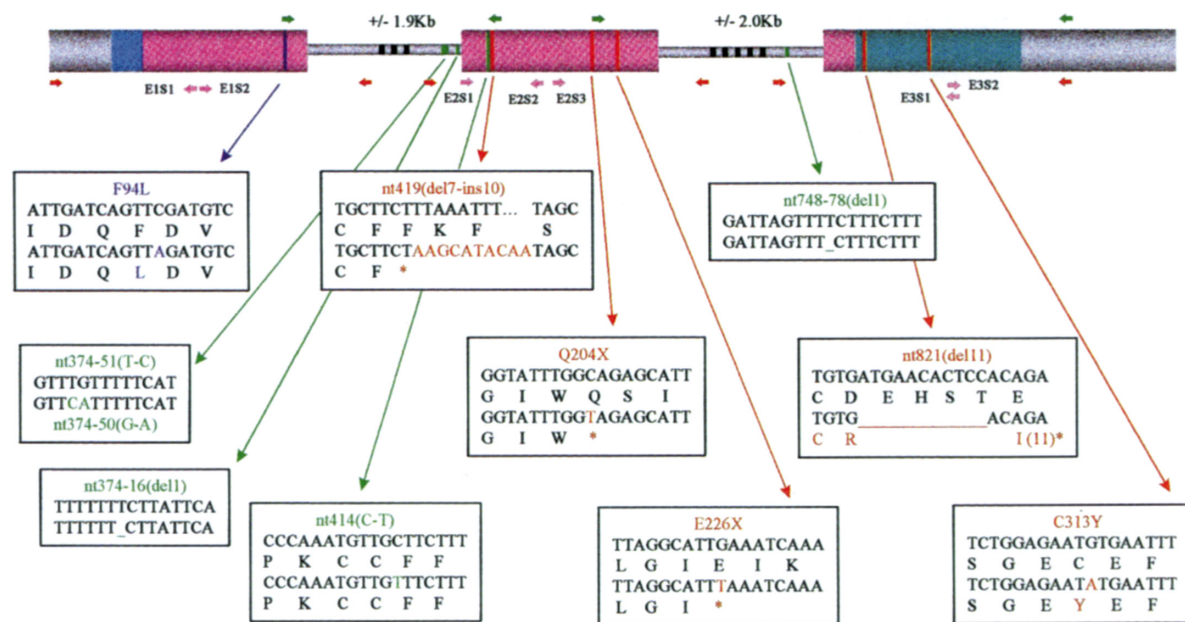
**Pedigree material.** In total, 32 animals with extreme muscular development were sampled from ten European beef cattle breeds in which double-muscled animals are acknowledged to occur at high to moderate frequency: (i) Belgium: Belgian Blue (4); (ii) France: Blonde d'Aquitaine (5), Charolais (2), Gasconne (2), Limousin (5), Maine-Anjou (4), Parthenaise (3); (iii) Spain: Asturiana (2), Rubia Gallega (2); (iv) Italy: Piedmontese (2). The determination of the double-muscled phenotype of the sampled animals was performed visually by experienced observers.

Four animals with conventional phenotype sampled from the Holstein-Friesian (2) and Jersey (2) dairy populations were analyzed as controls.

**Development of a primer set allowing for PCR amplification of the complete myostatin coding sequence from genomic DNA.** In order to facilitate the study of the myostatin coding sequence from genomic DNA, we determined the sequences of the exon-intron boundaries of the bovine gene. In mice, the myostatin gene is known to be interrupted by two introns, respectively ≈1.5 and 2.4 kb long (McPherron & Lee, 1997). We designed two primer pairs, respectively in bovine exons 1 and 2, exons 2 and 3, that were predicted to flank the two introns, assuming conservation of gene organization between mouse and cattle (Fig. 1 and Table 1). With

\* Both authors have contributed equally to this work.

Correspondence to: M. Georges



**Fig. 1.** Schematic representation of the bovine myostatin gene with position and definition of the identified DNA sequence polymorphisms. The gray boxes correspond to the untranslated leader and trailer sequences (large diameter), and the intronic sequences (small diameter) respectively. The blue, pink and green boxes correspond to the sequences coding for the leader peptide, N-terminal latency-associated peptide, and bioactive carboxyterminal domain of the protein respectively. Small green, red, and pink arrows point towards the positions of the primers used for intron amplification, exon amplification and sequencing, and exon sequencing

respectively; the corresponding primer sequences are reported in Table 1. The positions of the identified DNA sequence polymorphisms are shown as green, blue, or red lines on the myostatin gene for silent, conservative, and disrupting mutations respectively. Each mutation is connected via an arrow with a box reporting the details of the corresponding DNA sequence and eventually encoded peptide sequence. In each box, the variant sequence is compared with the control Holstein-Friesian sequence, and differences are highlighted in color.

**Table 1.** Primers used for PCR amplification and cycle sequencing.

Intron1-5'	5'-GAAGACGATGACTACACGCCAGGACG-3'	Intron1-3'	5'-CTAGTTTATTGTATTGTATCTTAGAGC-3'
Intron2-5'	5'-AGACTCCTACAACAGTGTGTTGT-3'	Intron2-3'	5'-ATACTCWAGGCCTAYAGCCTGTGGT-3'
Exon1-5'	5'-ATTCACTGGGTGGCAAGTTGTCTCTCAGA-3'	Exon1-3'	5'-CCCTCCTCCTTACATACAAGCCAGCAG-3'
Exon2-5'	5'-GTTTCATAGATTGATATGGAGGTGTTTCG-3'	Exon2-3'	5'-ATAAGCACAGGAACTGGTAGTTATT-3'
Exon3-5'	5'-GAAATGTGACATAAGCAAAATGATTAG-3'	Exon3-3'	5'-ATACTCWAGGCCTAYAGCCTGTGGT-3'
Exon1-Seq1	5'-TTGAGGATGTAGTGTGTTTCC-3'	Exon1-Seq2	5'-GCCATAAAAATCCAATCCTCAG-3'
Exon2-Seq1	5'-CATTTATAGCTGATCTTCTAACGCAAG-3'	Exon2-Seq2	5'-TGTCGAGGAGTCTTGACAGGCCCTCAG-3'
Exon2-Seq3	5'-GTACAAGGTACTGGAATCCGATCTC-3'		
Exon3-Seq1	5'-AGCAGGGGCCGGCTGAACCTCTGGG-3'	Exon3-Seq2	5'-CCCCAGAGGTTTCAGCCGCCCTGC-3'

these primer sets, two PCR products, respectively 2 kb and 3.5 kb long, were generated from a YAC clone (179A3) containing the bovine myostatin gene (Grobet et al., 1997). The PCR products were purified with QiaQuick PCR Purification kit (Qiagen) and partially sequenced with Dye terminator Cycle Sequencing Ready Reaction (Perkin Elmer) and an ABI373 automatic sequencer Alignment with the bovine cDNA sequence identified the four predicted exon-intron boundaries. Based on the available exonic and intronic sequences of the bovine myostatin gene, we subsequently designed three primer pairs that jointly allow for convenient amplification of the entire coding sequence from genomic DNA. The position of the corresponding primers is shown in Fig. 1, and the corresponding sequences are reported in Table 1.

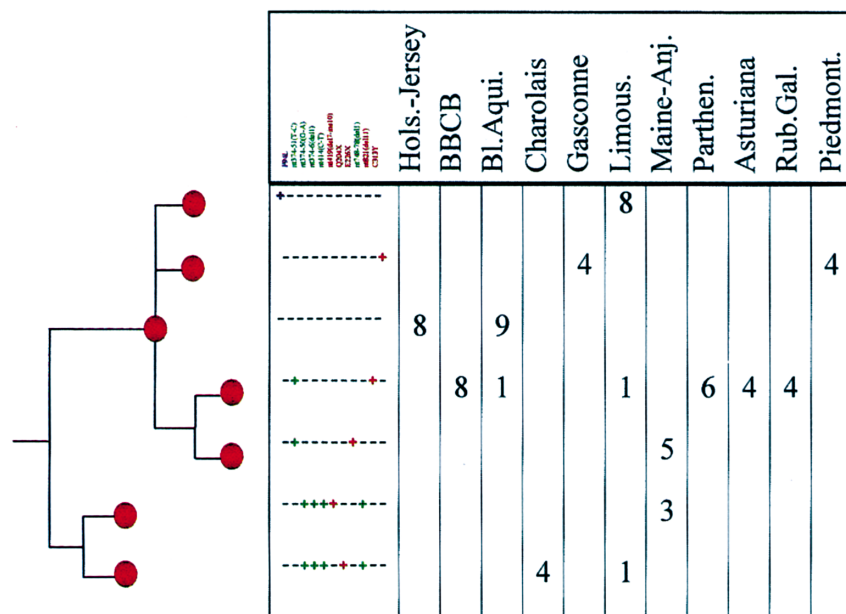
**Sequence determination of myostatin alleles.** After PCR amplification of the entire coding sequence from genomic DNA in the three described fragments, these were purified with QiaQuick PCR Purification kit (Qiagen) and sequenced with Dye terminator Cycle Sequencing Ready Reaction (Perkin Elmer) and an ABI373 automatic sequencer, using the primers used for amplification as well as a series of nested primers (Fig. 1 and Table 1). Chromat files produced with the ABI373 sequencer were analyzed with the Polyphred application (D. Nickerson, personal communication), which is part of a series of sequence analysis programs including Phred (B. Ewing, & P Green, unpublished). Phrap (P. Green, unpublished) and Consed (D. Gordon, unpublished).

## Results

**Identification of an allelic series of loss-of-function mutations in the bovine myostatin gene.** The coding sequence of the four control Holstein-Friesian and Jersey individuals was identical to the previously described wild-type allele (Grobet et al. 1997), indicating that we were indeed more than likely amplifying the genuine myostatin coding sequence and not dealing with a non-functional pseudogene.

Among the 32 double-muscled animals, on the contrary, we identified seven DNA sequence variants within the coding region, as summarized in Fig. 1.

In addition to the previously described *nt821(del11)* mutation in the third exon, deleting 11 base pairs at position 821 after the initiation codon and revealing a premature stop codon, we identified four new mutations that would be predicted to disrupt the myostatin function. An insertion/deletion at position 419 counting from the initiation codon, replacing 7 base pairs (bp) with an apparently unrelated stretch of 10 bp, reveals a premature stop codon in the N-terminal latency-associated peptide at amino acid position 140. This mutation is referred to as *nt419(del7-ins10)*. Two bp substitutions in the second exon, a C → T transition at



**Fig. 2.** Distribution of myostatin haplotypes by breed. The order and color of the myostatin mutations in the description of the observed haplotypes correspond to Fig. 1. All analyzed animals were double-muscled except for the two Holstein-Friesian and two Jerseys used as controls (column 1). A maximum parsimony cladogram connecting the different haplotypes is shown adjacent to the table. The rooting of the tree was based on the ovine myostatin sequence (data not shown).

nucleotide position 610 and a G → T transversion at nucleotide position 676, each yield a premature stop codon in the same N-terminal latency-associated peptide at amino-acid positions 204 and 226 respectively. These mutations are called *Q204X* and *E226X* respectively. Finally, a G → A transition at nucleotide position 938 results in the substitution of a cysteine by a tyrosine. This mutation is referred to as *C313Y*. This cysteine is the fifth of nine highly conserved cysteine residues typical of the members of the TGF-β superfamily and shared in particular by TGF-β1, -β2 and -β3, and inhibin-βA and -βB (McPherron and Lee 1996). It is thought to be involved in an intramolecular disulfide bridge stabilizing the three-dimensional conformation of the bioactive carboxyterminal peptide. Its substitution is, therefore, likely to affect the structure and function of the protein. This *C313Y* has recently also been described by Kambadur and associates (1997).

We also identified a conservative phenylalanine-to-leucine substitution at amino acid position 94 in the first exon, owing to a C → A transversion at nucleotide position 282 of the myostatin gene. Given the conservative nature of the amino acid substitution, its location in the less conserved N-terminal latency-associated peptide, and as this mutation was observed at the homozygous condition in animals that were not showing any sign of exceptional muscular development, we predict that this mutation does not interfere drastically with the myostatic function of the encoded protein, if at all. This mutation is referred to as *F94L*. Note that the murine protein is characterized by a tyrosine at the corresponding amino acid position.

Finally, we identified a silent C → T transition at the third position of the 138th cytosine codon in the second exon, referred to as *nt414(C-T)*.

In addition to these DNA sequence polymorphisms detected in the coding region of the myostatin gene, we identified four DNA sequence variants in intronic sequences that are probably neutral polymorphisms and that have been assigned the following symbols: *nt374-51(T-C)*, *nt374-50(G-A)*, *nt374-16(del1)* in intron 1, and *nt748-78(del1)* in intron 2 (Fig. 1).

**Distribution of myostatin haplotypes in different cattle breeds.** Figure 2 reports the observed myostatin haplotypes as well as their distribution in the analyzed sample sorted by breed.

For the majority of the studied breeds, the analyzed double-

muscled animals were homozygous for one of the five described mutations predicted to disrupt the myostatin function or compound heterozygotes for two distinct of these mutations. This is compatible with the hypothesis that the double-muscled condition has a recessive mode of inheritance in all these breeds.

Only in Limousin and Blonde d'Aquitaine could we not demonstrate clear evidence for the role of myostatin loss-of-function mutations in the determinism of the observed muscular hypertrophy. Most Limousin animals were homozygous for the conservative *F94L* substitution which is unlikely to cause the muscular hypertrophy characterizing these animals for the reasons cited above. Two Limousin animals proved to be heterozygous for this mutation, the other allele being either the *nt821(del11)* or *Q204X* mutation. All Blonde d'Aquitaine animals but one were homozygous wild-type, the latter carrying the *nt821(del11)* mutation. These data indicate either that the myostatin gene is not involved in the double-muscled condition characterizing these two breeds, or that there are additional myostatin mutations outside of the coding region. It should be noted that the double-muscling condition is often considered to be less pronounced in Limousin animals than in other breeds; this may point towards a phenocopy with a distinct determinism in this breed.

Although the analyzed sample is very limited so far, our data indicate that some mutations, such as the *nt821(del11)* and *C313Y*, are shared by several breeds, which points towards gene migration between the corresponding populations, while others seem to be confined to specific breeds. Moreover, while some breeds (the Belgian Blue breed in particular) seem to be essentially genetically homogeneous, others show clear evidence for allelic heterogeneity (for example, Maine-Anjou).

Figure 2 also depicts a preliminary maximum parsimony cladogram predicting the evolutionary relationships between the identified haplotypes. The tree was rooted based on the observation that the ovine myostatin sequence exhibited the *nt374-50(G-A)*, *nt374-16(del1)* and *nt748-78(del1)* variants. It can be seen from this figure that the *nt419(del7-ins10)* and *Q204X* mutations, encountered respectively in the French breeds Maine-Anjou and Charolais/Limousin, appear to have occurred in the same "ancestral" myostatin haplotype characterized by the *nt374-50(G-A)*, *nt374-16(del1)*, *nt748-78(del1)*, and *nt414(C-T)* neutral variants (the former three being shared with the ovine haplotype). Like-

wise, the *E226X* and *nt821(del11)* mutations seem to have occurred in a common, more recently evolved haplotype, characterized by the *nt374-51(T-C)* variant.

## Discussion

We demonstrate in this work that the double-muscling phenotype is genetically heterogeneous in cattle, involving at least five different mutations in the bovine myostatin gene.

The observation of allelic heterogeneity contradicts the classical view that a single *mh* mutation spread through the European continent in the beginning of the 19th century with the dissemination of the Shorthorn breed from the British Isles (Méniissier, 1984). Two of the mutations, at least, are shared by more than one breed, indicating some degree of gene migration, but definitely not from a single origin.

While the observed level of allelic heterogeneity might reflect a particularly high mutation rate for this gene, we are inclined to believe that it results from the long-standing history of artificial selection for meaty animals, which must have favored such mutations and maintained them in the corresponding populations. In the BBB, evidence has been gathered in favor of a modest but significant superiority of muscular development in heterozygous (*nt821(del11)/+*) over homozygous (*+/+*) normal individuals, reflecting some degree of haploinsufficiency (Hanset and Michaux, 1985a, 1985b). A similar tendency was apparent from the weight and carcass analysis of myostatin knock-out mice (McPherron et al. 1997). Even a modest expression in heterozygotes may considerably have increased the selective advantage of the bovine myostatin loss-of-function alleles and therefore their population frequency.

The observation that in at least eight of the ten studied breeds double-muscling involves five independent myostatin mutations indicates that the number of genes susceptible to affect muscular development in a comparable way is likely to be limited in cattle. It is tempting, therefore, to anticipate that myostatin similarly plays a key role in other species as well. To the best of our knowledge, however, no evidence is available to date of muscular hypertrophies involving this gene in livestock species other than cattle. The exceptional muscularity characterizing callipyge sheep involves a gene located on sheep Chr 18, which is orthologous to bovine Chr 21 (Cockett et al. 1994, 1996). The meatiness typical of Piétrain pigs is determined, at least to some degree, by the *CRC* or a tightly linked gene on pig Chr 6 (Fuji et al. 1991). The *CRC* gene has been mapped to Chr BTA18q23-24 in cattle. In contradiction to double-muscling in cattle, which from an histological point of view is a hyperplasia, the enhanced muscular development of callipyge sheep and Piétrain pigs involves a true hypertrophy of the muscle fibers, that is, an increase in their individual diameter.

In mice, in addition to the in vitro generated myostatin knock-out mice (McPherron and Lee 1997), the compact mutation could be due to a naturally occurring mutation at the myostatin gene. The compact locus has indeed been mapped to the *D1Mit375-D1Mit21* interval on mouse Chr 1 known to be orthologous to *HSA2q31-32* and *BTA2q12-22* (Varga et al. 1997).

From an applied point of view, the characterization of a panel of mutations in the myostatin gene associated with double-muscling will contribute to the establishment of a diagnostic

screening system allowing for marker-assisted selection for or against this condition in cattle. The observed high level of genetic heterogeneity of double-muscling forces one to carefully evaluate the relevance of the already identified mutations in each population and eventually to seek other causative mutations.

**Acknowledgments.** This work was supported by a grant from the Belgian Ministère de l' Agriculture et des Classes Moyennes (D1/2-5744A). LJR was funded by a grant from the CYCIT(GAN95-0853). We are very grateful to Deborah Nickerson, David Gordon, and Phil Green for providing us the Polyphred application. We sincerely thank Professor Pascal Leroy for his continuous support and interest in our work.

## References

- Charlier C, Coppeters W, Farnir F, Grobet L, Leroy P, Michaux C, Mni M, Schwens A, Vanmanshoven P, Hanset R, Georges M (1995) The *mh* gene causing double-muscling in cattle maps to bovine chromosome 2. *Mammalian Genome*, 6, 788–792
- Cockett NE, Jackson SP, Shay TD, Nielsen D, Green RD, Georges M (1994) Chromosomal localization of the callipyge gene in sheep (*Ovis aries*) using bovine DNA markers. *Proc Nat Acad Sci USA* 91, 3019–3023
- Cockett NE, Jackson SP, Shay TD, Farnir F, Berghmans S, Snowder G, Nielsen D, Georges M (1996) Polar overdominance at the ovine callipyge locus. *Science* 273, 236–238
- Culley G (1807) *Observations of Livestock*, 4th ed. (London: G. Woodfall)
- Dunner S, Charlier C, Farnir F, Brouwers B, Canon J, Georges M (1997) Towards interbreed IBD fine mapping of the *mh* locus: double-muscling in the *Asturiana de los Valles* breed involves the same locus as in the *Belgian Blue* cattle breed. *Mamm Genome* 8, 430–435
- Fuji J, Otsu K, Zorzato F, Deleon S, Khanna VK, Weiler JE, O'Brien PJ, MacLennan DH (1991) Identification of a mutation in the porcine ryanodine receptor associated with malignant hyperthermia. *Science* 253, 448–451
- Georges M, Andersson L (1996) Livestock genomics comes of age. *Genome Res*, in press
- Grobet L, Royo Martin LJ, Poncelet D, Pirotin D, Brouwers B, Riquet J, Schoeberlein A, Dunner S, Menissier F, Massabanda J, Fries R, Hanset R, Georges M (1997) A deletion in the myostatin gene causes double-muscling in cattle. *Nature Genet* 17, 71–74
- Hanset R (1991) The major gene of muscular hypertrophy in the Belgian Blue Cattle breed. In *Breeding for Disease Resistance in Farm Animals*, JB Owen, RFE Axford (eds.) ( : CAB International), pp 476–478
- Hanset R, Michaux C (1985a) On the genetic determinism of muscular hypertrophy in the Belgian White and Blue cattle breed. I. Experimental data. *Genet Sel Evol (Paris)* 17, 359–368
- Hanset R, Michaux C (1985b) On the genetic determinism of muscular hypertrophy in the Belgian White and Blue cattle breed. II. Population data. *Genet Sel Evol (Paris)* 17:369–386
- Kambadur R, Sharma M, Smith TPL, Bass JJ (1997) Mutations in myostatin (GDF8) in double-muscled Belgian Blue Cattle. *Genome Res* 7, 910–916
- McPherron AC, Lee S-J (1996) The transforming growth factor  $\beta$  superfamily. In *Growth Factors and Cytokines in Health and Disease*, Vol 1B. ( : JAI Press Inc.), pp 357–393
- McPherron AC, Lawler AM, Lee S-J (1997) Regulation of skeletal muscle mass in mice by a new TGF $\beta$  superfamily member. *Nature* 387, 83–90
- Méniissier F (1982) Present state of knowledge about the genetic determination of muscular hypertrophy or the double muscled trait in cattle. In *Curr Top Vet Med Anim Sci* 16, 387–428
- Varga L, Szabo G, Darvasi A, Müller G, Sass M, Soller M (1997) Inheritance and mapping of compact (*Cmpt*), a new mutation causing hypermuscularity in mice. *Genetics*, in press

# High-resolution, human–bovine comparative mapping based on a closed YAC contig spanning the bovine *mh* locus

Dimitri Pirottin,<sup>1,\*</sup> Dominique Poncelet,<sup>1,\*</sup> Luc Grobet,<sup>1</sup> Luis José Royo,<sup>2</sup> Benoit Brouwers,<sup>1</sup> Julio Masabanda,<sup>3</sup> Haruko Takeda,<sup>4</sup> Ruedi Fries,<sup>3</sup> Yoshikazu Sugimoto,<sup>4</sup> James E. Womack,<sup>5</sup> Susana Dunner,<sup>2</sup> Michel Georges<sup>1</sup>

<sup>1</sup>Department of Genetics, Faculty of Veterinary Medicine, University of Liège (B43), 20 Bd de Colonster, 4000-Liège, Belgium

<sup>2</sup>Laboratorio de Genética, Dpto. de Producción Animal, Universidad Complutense, 28040 Madrid, Spain

<sup>3</sup>Technische Universität München (TUM), D-85350 Freising-Weihenstephan, Germany

<sup>4</sup>Shirakawa Institute of Animal Genetics, Japan

<sup>5</sup>Department of Veterinary Pathobiology, Texas A&M, College Station, Texas, USA

Received: 6 August 1998 / Accepted: 28 October 1998

**Abstract.** A closed YAC contig spanning the *mh* locus was assembled by STS content mapping with seven microsatellite markers, eight genes or EST, and nine STS corresponding to YAC ends. The contig comprises 27 YACs, has an average depth of 4.3 YACs, and spans an estimated 1.2 Mb. A linkage map was constructed based on five of the microsatellite markers anchored to the contig and shown to span 7 cM, yielding a ratio of 160 kb/1 cM for the corresponding chromosome region. Comparative mapping data indicate that the constructed contig spans an evolutionary breakpoint connecting two chromosome segments that are syntenic but not adjacent in the human. Consolidation of human gene order by means of whole genome radiation hybrids and its comparison with the bovine order as inferred from the contig confirm conservation of gene order within segments.

## Introduction

In recent years, we have witnessed the multiplication of efforts to map economic trait loci (ETL) in livestock with strategies based on linkage analysis (for example, Georges 1998). Although mapping data can be exploited via Marker Assisted Selection (Visscher et al. 1998), maximal valorization probably requires actual cloning of the causal genes (Andersson 1998). When attempting to positionally clone ETL in livestock, animal geneticists benefit to a great extent from the remarkable progress achieved in characterizing the human and mouse genomes. Indeed, conserved chromosome segments (that is, segments within which gene order is conserved for two or more species) are estimated to be of the order of 10 cM on average between human and mouse (Copeland et al. 1993) and possibly larger between man and most livestock species. Identifying these conserved segments and their boundaries in the region of interest therefore allows for subsequent “comparative positional candidate cloning” (Archibald 1998) by exploiting human or mouse transcript maps.

Inference about the extent of gene order conservation is, however, affected by the resolution of the applied mapping methods. Indeed, refined mapping methods are susceptible to reveal rearrangements that could not be resolved by coarser methods. To verify whether micro-rearrangements might be common within previously defined conserved segments, we have generated a high-resolution, bovine–human comparative map surrounding the myostatin (*MSTN*) locus.

Recently, independent efforts led to the demonstration that

loss-of-function mutations in the *MSTN* gene cause double-muscling in mouse and cattle (McPherron et al. 1997; Grobet et al. 1997, 1998; Kambadur et al. 1997; McPherron and Lee 1997). One of these studies used a comparative positional candidate strategy to identify the culprit bovine gene (Grobet et al. 1997). In this study, the bovine BTA2q1 linkage map and human HSA2q31-32 transcript RH-map were aligned using coincident bovine YACs, that is, large insert clones that contained both a bovine microsatellite marker and the bovine ortholog of a mapped human EST. This led to the identification of an interval on the human transcript RH-map that was likely to contain the double-muscling gene. The demonstration that *MSTN*, known to cause a muscular hyperplasia when knocked out in mice, mapped to that very interval, and the subsequent identification of loss-of-function mutations among double-musclered individuals demonstrated the role of *MSTN* in the determination of the double-muscling phenotype.

In this paper, we describe the completion of a YAC-based contig spanning the bovine *MSTN* locus. Bovine microsatellites anchored to this contig allow for a comparison of the linkage and physical map of the region. ESTs mapping to the corresponding region on the human map have been positioned on the bovine contig, allowing for a high-resolution comparison of both maps.

## Materials and methods

**YAC contig construction.** Two distinct bovine YAC libraries representing a total of 12 genome equivalents were used in this work (Schoeberlein et al. unpublished; Takeda et al. 1998). PCR screening of the YAC libraries was performed with three dimensional pooling schemes as described (Libert et al. 1993; Takeda et al. 1998). High-density filters of the Schoeberlein YAC library were generated and used to screen the library by hybridization according to Cai and associates (1996). YAC ends were isolated using “bubble-PCR” following Libert and colleagues (1993). Microsatellite markers were isolated from YACs according to Cornelis and coworkers (1992). Map position of the YAC ends and microsatellites was examined by genotyping either a panel of bovine–rodent somatic cell hybrids (Dietz et al. 1992) or a panel of bovine–rodent radiation hybrids (Womack et al. 1997). Primer pairs used for PCR amplification of the sequence-tagged sites (STS) utilized for contig construction are as described (Grobet et al. 1997; Sonstegard et al. 1997) or reported in Table 1. FISH hybridization with total YAC clone DNA as probes was performed as previously described (Solinas-Toldo et al. 1995b; Masabanda et al. 1998).

The OPTICONTIG program (Georges unpublished) was used to determine the most likely STS/probe order based on STS/probe content of individual YACs. Starting from a randomized order, OPTICONTIG generates a Markov chain in which the  $n + 1$  state differs from the  $n$  state by the random reassignment of the position of one randomly selected marker. For each state, a criterion is computed counting the number of negative STS within YAC boundaries given the considered order, and penalizing

\* Both authors contributed equally to this work.

Correspondence to: M. Georges



**Table 1.** Primers pairs used for PCR amplification of YAC ends and genes.

YAC ends		
117C7-T	TTTAACCTCTGAGCCACCGGAAGC	CAGCTATTAGTAGCATTCTGAGGT
117C7-U	TAATAGGTGTTTGTAAAAGCAAATC	ATCAATATAAAAATGGGCTAAATGCGG
74E11-T	GCATCAATATAAAAATGGGCTAAATGCGG	GTAAGAAGCAAATCCACTGTTTTGGGGGAATG
215H8-T	AAGACAGAGTAGGTGGGAATGGGT	ATTACAGGATAAGGTCAGATCCC
215H8-U	AAGTAAGAACTAGATAAACTGTG	GTCTCTTGGACTTTTTCAGGTATATG
189A3-U	GACTATATAAATTGTAACTTCAACAATTG	ATCCGTAACAAATTACTTGGGCGTGGCGTC
179A3-T	CACCTCCACTAGGTGTGTTCTGATG	AGGAATGCAAGCCTACTTTAATACC
179A3-U	GCCTGGAAGGATGGCACTGCCAGCTTT	CAAAGCTGGCAGTGCCATCCTCCAGG
178C6-U	CTTTGATCATTTCCTCAATTTATGTGC	CAGAGGACCCTGAAGGGTTATAGTGC
Genes		
NAB1	GTAAACCCATCCAGAGTAATG	TTGATGCCTTTATCCAACAAGGTGG

orders which would not place STS corresponding to YAC ends at the extremities of the corresponding YAC. OPTICONTIG utilizes simulated annealing (for example, Weir 1996) to identify the order minimizing the value of the criterion. The chain is run several times to monitor repeated convergence to the same optimal order or set of equally optimal orders.

**Genetic map construction.** Microsatellite genotyping was performed as previously described (Georges et al. 1995). Linkage analyses were performed with the TWOPT and BUILD options of the CRIMAP package (Lander and Green 1987). Double recombinant individuals were identified with the CHROMPIC option and regenotyped.

**Construction of a radiation hybrid map.** DNA from the Genebridge-4 (Walter et al. 1994) and Stanford TNG whole genome radiation hybrid panels was obtained from Research Genetics (Huntsville, Ala.). Both panels were genotyped for human STS with primer pairs whose sequences were obtained from the Whitehead Institute/MIT Centre for Genome Research (Hudson et al. 1995). The resulting segregation vectors were analyzed with the RHMAXLIK program from the RHMAP package (Lunetta et al. 1995). The analysis utilized information from both panels simultaneously, using non-proportional distances between markers among the panels.

## Results

**Construction of a YAC contig spanning the bovine *mh* locus.** In a previous study (Grobet et al. 1997), we described the construction of two small YAC contigs postulated to flank the *mh* locus. Contig "A", comprising seven YACs, contained microsatellite markers *TGLA44*, *BM81124*, *BULGE23*, and *BULGE27*, and the *INPPI* gene. Contig "B", comprising five YACs, contained microsatellites *BULGE20*, *BULGE28*, and *Col3A1*. The *MSTN* gene, causing the double muscling phenotype in cattle and mice, was predicted to map between these two contigs based on genetic mapping of the *mh* locus as well as on comparative mapping data (see above). One YAC containing *MSTN* (Y179A3) was isolated in the same study but could not be connected with either contig "A" nor "B". Therefore, the actual physical distance separating both contigs remained unknown, and the precise position of the myostatin gene on the bovine map was not formally demonstrated.

To complete the physical map of this genomic region, we therefore screened a six-genome equivalent bovine YAC library (Schoeberlein et al. unpublished) by PCR with two previously described microsatellite markers isolated respectively from contig "A" (*BULGE27*) and contig "B" (*BULGE20*), as well as six novel sequence-tagged sites (STS). Five of these were developed from YAC ends isolated by bubble-PCR from clones reported in Grobet et al. (1997) (117C7-T, 117C7-U, 215H8-T, 179A3-T, and 178C6-U). The remaining one corresponded to an STS amplifying the bovine ortholog of a gene known to map in the vicinity of the *INPPI*-*Col3A1* interval on the human radiation hybrid map: the

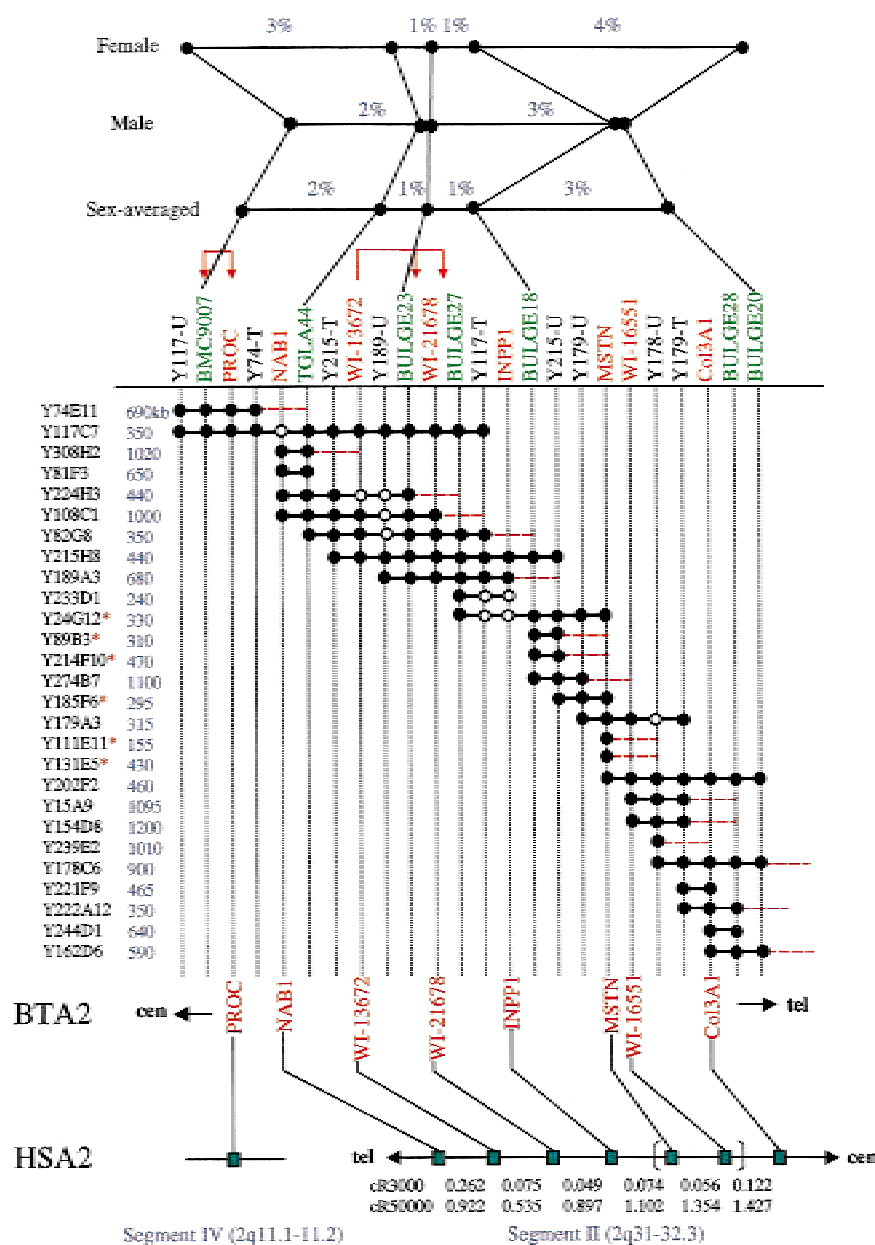
human NAB1 transcriptional repressor (NAB1; Hudson et al. 1995). The corresponding primers were designed by aligning the human and rat NAB1 sequence and targeting primer sequences towards highly conserved segments of the gene. The same YAC library was also screened by filter hybridization with pooled cDNA clones (Image id.: 149056 and 321388) corresponding to an expressed sequence tag (EST) mapping to the same region on the human map: WI-16551. A second bovine YAC library (Takeda et al. 1998) was subsequently screened by PCR with two additional YAC ends (215H8-U and 179A3-U) as well as with *MSTN*.

Altogether, we isolated 14 new YACs, yielding a total of 27 YACs in the region. FISH hybridization was used to confirm the localization of 24 of these YACs to the centromeric end of bovine Chromosome (Chr) 2 and identify possible chimerism. All selected YACs yielded a signal at the expected chromosome position, while 17 YACs generated additional signals on one (8) or more (9) other chromosomes. Pulsed field gel electrophoresis followed by Southern blot hybridization to total genomic DNA allowed us to estimate the individual size of each YAC, yielding an average insert size of 548 kb. Figure 1 reports the estimated size for all isolated YACs.

The presence of all STS and probes that were utilized to screen the libraries was confirmed on preparations of individual DNA. Moreover, the 27 YAC DNAs were tested by PCR for the presence of (i) two additional microsatellite markers: *BMC9007* (*PROC*; Sonstegaard et al. 1997) and a new microsatellite isolated from YAC 215H8: *BULGE18*; (ii) two novel YAC ends isolated by bubble-PCR: 74E11-T and 189A3-U; and (iii) an STS amplifying bovine protein C (*PROC*; Sonstegaard et al. 1997). In addition, the presence of the bovine orthologs of two additional EST (WI-13672, WI-21678) mapping in the vicinity of the *INPPI*-*Col3A1* interval in human (Hudson et al. 1995), was tested on individual YAC DNA by Southern blotting with human cDNA clones (respectively Image clones 33009 and 172303) as probes under low-stringency hybridization conditions.

The STS/probe content allowed us to group all 27 YACs into a single, closed contig as shown in Fig. 1. The STS/probe order minimizing the number of inconsistencies in the YAC contig was identified with the OPTICONTIG program. Six equally parsimonious orders were identified and are reported in Fig. 1. These results form the possibility to flip the adjacent *PROC* and *BMC9007* markers, as expected given the fact that both STS were isolated from the same cosmid (Sonstegaard et al. 1997); as well as three equally parsimonious positions of EST WI-13672. Six of the 27 YACs were found positive for *MSTN*, therefore clearly demonstrating that *MSTN* indeed maps to this contig. Moreover, STS content data positioned *MSTN* between *INPPI* and *Col3A1* as predicted from the human transcript map (Grobet et al. 1997).

Analysis of Fig. 1 shows that three YACs (117C7, 24G12, and 202F2) are sufficient to define a minimum tiling path spanning the entire contig. Based on the estimated size of these clones (117C7:



**Fig. 1.** YAC contig spanning the *MSTN* locus in cattle. STS/probes are represented in the most parsimonious order as determined from the STS content of individual YACs. Genes are represented in red, microsatellites in green, and YAC ends in black. Arrows point towards alternative, equally parsimonious positions for some of the STS. Black bullets report STS content of individual YACs while white bullets represent conflicts in the contig, given marker order. YACs isolated from the Takeda and coworkers (1998) library are marked with a red asterisk. Sizes of individual YACs in kilobases are reported in blue adjacent to YAC names. Chimeric YACs as determined by FISH hybridization are marked with a red stippled line. Sex-averaged, as well as male- and female-specific recombination rates between five microsatellite markers anchored to the contig, are shown above the contig. The gene order obtained in cattle is compared with the gene order determined with radiation hybrids in human at the bottom of the figure. The distance between adjacent genes on the human map is given in centirays as estimated with the Genebridge-4 and TNGF panels. Markers that could not be ordered with odds >100:1 are bounded by red brackets. Genes are assigned to evolutionary conserved chromosome segments II and IV as defined in Sonstegaard and associates (1997). The positions of the centromere and telomere with respect to the gene array show the inversion of segment II when comparing human and bovine.

350 kb; 24G12: 330 kb; 202F2: 460 kb), the corresponding chromosome region might be as small as 1140 kb. It is noteworthy that YAC 24G12 is chimeric, which would lead to an overestimation of the actual physical size of the region. On the other hand, the YACs could carry deletions, as indicated by some inconsistencies in the STS content, which would lead to an underestimation of the total size.

**Construction of a high-resolution genetic map spanning the *mh* locus.** The seven microsatellites included in the YAC contig were utilized to genotype 331 individuals from the bovine three-generation IBRP reference family (Hetzl et al. 1993), 255 individuals from the previously described Sart-Tilman backcross pedigree (Charlier et al. 1995), as well as 96 individuals corresponding to five paternal half-sib Maine-Anjou pedigrees segregating for the double-muscling trait (Grobet et al. unpublished). Five of the seven microsatellite markers proved to be polymorphic in this pedigree material.

The corresponding genotypes were used to construct a linkage map with CRIMAP (Lander and Green 1987). The most likely order and sex-averaged recombination rates between adjacent markers are represented in Fig. 1. Odds versus alternative orders were >1000. It can be seen that the marker order resulting from the linkage analysis coincided with the order as determined from the YAC contig. The distance between the outermost markers, *BMC9007* (*PROC*) and *BULGE20*, was estimated at 7 cM. Therefore, this suggests a ratio of the order of 160 kb per cM in this chromosome region.

When allowing for sex-specific recombination rates, we obtained the male and female maps shown in Fig. 1. It can be seen that for three of the four marker intervals, the estimated male recombination rates are smaller than the female estimates. For the *BULGE23*–*BULGE18* interval, on the contrary, the recombination rate seems larger in the heterogametic meioses. The likelihood of the data under the model of sex-specific recombination rates is, however, only 25 times superior when compared with the sex-

averaged model. More data are, therefore, needed to assess the significance of these sex-specific differences in recombination rates.

**High-resolution human/bovine comparative mapping.** Zoo-FISH experiments have shown that along most of its length bovine Chr 2 (BTA2q12-42) harbors sequences whose orthologs map to the long arm of Chr 2 in the human. Only the distal end of bovine Chr 2 (BTA2q43-45) seems to escape this rule, the orthologous sequences mapping to distal HSA1p in the human (Solinas-Toldo et al. 1995a). Subsequent work demonstrated that most of BTA2 corresponds to a complex rearrangement of five HSA2q segments (Sonstegard et al. 1997). The precise evolutionary breakpoints as well as the relative orientation of these segments in human and bovine, however, remain poorly defined.

From the genes assigned to the contig described in this work, *PROC* maps to 2q13-21.3 in the human (segment IV in Sonstegard et al. 1997), while the seven other ones (*NAB1*, *WI13672*, *WI21678*, *INPP1*, *MSTN*, *WI16551* and *COL3A1*) are mapping to 2q31-32.3 (segment II in Sonstegard et al. 1997). Therefore, the most proximal YACs from our contig (particularly *Y117C7*) are bound to contain the corresponding evolutionary breakpoint (Fig. 1).

To more precisely assess the relative orientation of segment II in human and cattle, as well as to determine whether additional micro-rearrangements might have occurred within segment II during evolution, we re-examined the precise order of the seven corresponding genes on the human map. Indeed, while the approximate location of these genes on the human transcript map is known, their ordering based on available data remains ambiguous. We therefore genotyped both a 3000 rad (Genebridge 4; Walter et al. 1994) and a 50,000 rad (Stanford TNG RH panel) human whole genome hybrid panel for the corresponding genes. The resulting segregation vectors were analyzed jointly with the RHMAXLIK option of the RHMAP package (Lunetta et al. 1995). As illustrated in Fig. 1, this analysis resulted in reliable ordering of all but two genes (*MSTN* and *WI16551*), whose position could be flipped without significantly affecting the likelihood of the data. All other orders could be rejected based on odds >100. Therefore, these data point towards a conservation of gene order within the analyzed chromosome microsegment. The orientation of the gene cluster with respect to flanking framework markers included both in linkage maps and in the radiation hybrid maps (Hudson et al. 1995) indicates the inversion of segment II in the bovine with respect to the centromere (Fig. 1).

## Discussion

We report the construction of a YAC-based contig spanning the *TGLA44-BULGE20* interval on proximal BTA2q. By assigning the *MSTN* gene to this contig, we confirm the location of the *mh* locus causing double-muscling in cattle to this marker interval as anticipated from linkage studies and human-bovine comparative mapping data (Grobet et al. 1997, Smith et al. 1997).

By comparing estimates of the physical distance (base pairs) covered by this contig with the genetic distance (centimorgan) separating the outermost informative microsatellite markers anchored to the contig (*BMC9007* and *BULGE20*), we calculated a ratio of 160 kb per cM for this chromosome region. This estimate has to be considered cautiously given the evidence for a high incidence of chimerism characterizing the utilized YAC libraries and the possibility for internal deletions affecting some of the YAC clones. Nevertheless, the obtained ratio is remarkably low. Indeed, by analogy with other mammals, the haploid DNA content of the bovine genome is usually estimated at approximately three billion base pairs, while the total sex-averaged genetic map length has been estimated at approximately 30 Morgan (Kappes et al.

1997). Therefore, one would expect a ratio of one million base pairs per centimorgan on average, far superior to what is suggested by our data for the studied chromosome region. In human, recombination rates are known to be inflated towards the telomeric end of chromosome arms, while the opposite tendency is found around centromeres. Our results suggest an inflation of recombination rates in a sub-centromeric region. Note that in bovine all chromosomes are acrocentric.

This work provides us with two independent estimates of the degree of chimerism characterizing the utilized YAC library (Schoeberlein et al. unpublished). Clone extremities were isolated from a number of YACs with bubble-PCR, and their map positions were determined with panels of either bovine-rodent somatic cell hybrids (Dietz et al. 1992) or bovine-rodent radiation hybrids (Womack et al. 1997). YAC end cloning was performed prior to receiving the outcome from the FISH mapping experiments, therefore without bias with regard to possible chimerism. Based on the corresponding results, YACs were sorted in three classes: (i) three nonchimeric YACs for which both extremities were mapping to BTA2q1; (ii) four chimeric YACs for which either both or one extremity was mapped to a chromosome region other than BTA2q1; and (iii) three ambiguous YACs for which one YAC end was shown to map to the right chromosome region, while no data were available for the other extremity. Expressing the likelihood of these data as a simple function of the degree of chimerism,  $C$ ,

$$L = (1 - C)^3 C^4 (1 - 0.5C)^3$$

yields a maximum likelihood estimate of  $C = 52\%$ . FISH mapping, on the other hand, yielded an estimate of 17/24, or 71% chimerism. While both estimates may not be significantly different given the limited sample size, the estimate from the FISH hybridization seems unusually high. Examination of the distribution of the signals on the chromosomes other than BTA2q1, however, revealed evidence of non-randomness as five from the 24 recorded signals mapped to 1q31-35, while four mapped to 5q19-25. Although these results have to be considered cautiously, these data might indicate the existence of paralogous loci in the bovine genome. Similar results have recently been reported for the human genome (Eichler et al. 1998).

A previously described 5000 rad whole genome hamster-bovine radiation hybrid panel (Womack et al. 1997) proved to be very valuable to determine whether newly developed STS mapped to the predicted chromosome region. This resource could, however, not be used in this work to assist in ordering STS within the region spanned by the constructed contig owing to a lack of reliable chromosome breaks in this region. This is not unexpected given the limited physical size covered.

Detailed analysis of the map order of seven genes mapping to a conserved chromosome segment in human and cattle (segment II in Sonstegard et al. 1997) indicates interspecies conservation of gene order within this segment. This finding supports the generally admitted assumption that chromosome evolution during mammalian radiation has been accompanied by reshuffling of conserved chromosome segments; however, that within such segments gene order would be faithfully conserved. When comparing human and mice, for instance, the length of these conserved segments has been estimated at 8.8 cM (Copeland et al. 1993). This information is important for future positional candidate cloning efforts in livestock species, as it asserts that once the boundaries of conserved chromosome segments are identified, the order on the human transcript map accurately predicts gene order in the studied animal species.

A microsatellite (*BMC9007*) isolated from a bovine cosmid containing the *PROC* gene was shown to map to the assembled contig. *PROC* is known to map to HSA2q13-21.3 on the human map, which corresponds to conserved chromosome segment IV as defined by Sonstegard and coworkers (1997). Therefore, our con-

tig must contain the evolutionary breakpoint linking segments II and IV in the bovine. This will eventually facilitate its cloning and molecular characterization.

**Acknowledgments.** This work was supported by a grant from the Belgian Ministère de l'Agriculture et des Classes Moyennes and Skye PharmaTech (Toronto) (D1/2-5744A) and a grant from the European Union (B104-CT95-0073). L.J. Royo was funded by a grant from the CYCIT (GAN95-0853). Construction of the bovine RH-panel was funded by grant No. 97-00675 from the USDA-NRI to Professor J.E. Womack. We sincerely thank Dr Juliette Riquet, Dr Carine Nezer, and Professor Pascal Leroy for stimulating discussions.

## References

- Andersson L (1998) Identification and cloning of trait genes. In *Animal Breeding: Technology for the 21st Century*, J Clark, ed. (Amsterdam: Harwood Academic Publishers), pp 103–118
- Archibald AL (1998) Comparative genome mapping—the livestock perspective. In *Animal Breeding: Technology for the 21st Century*, J Clark, ed. (Amsterdam: Harwood Academic Publishers), pp 137–164
- Cai L, Schalkwijk LC, Stehli AS, Zee RYL, Haaf T et al. (1996) Construction and characterization of a 10-genome equivalent yeast artificial chromosome library for the laboratory rat, *Rattus norvegicus*. *Genomics* 38, 385–392
- Charlier C, Coppieters W, Farnir F, Grobet L, Leroy P et al. (1995) The mh gene causing double-muscling in cattle maps to bovine Chromosome 2. *Mamm Genome* 6, 788–792
- Copeland NG, Jenkins NA, Gilbert DJ, Eppig JT, Maltais LJ et al. (1993) A genetic linkage map of the mouse: current applications and future prospects. *Science* 262, 57–66
- Cornelis F, Hashimoto L, Loveridge J, MacCarthy A, Buckle V et al. (1992) Identification of a CA repeat at the TCRA locus using YACs: a general method for generating highly polymorphic markers at chosen loci. *Genomics* 13, 820–825
- Dietz AB, Georges M, Threadgill DW, Womack JE, Schuler LA (1992) Somatic cell mapping, polymorphism, and linkage analysis of bovine prolactin-related proteins and placental lactogen. *Genomics* 14, 137–143
- Eichler EE, Adams M, Schwartz S, Rocchi M, Mohrenweider HW (1998) Pericentromeric repeats and the expansion of gene families in the human genome. *Proceedings of the 1998 Cold Spring Laboratory Meeting on Genome Mapping, Sequencing & Biology*, May 13–May 17
- Georges M (1998) Mapping genes underlying production traits in livestock. In *Animal Breeding: Technology for the 21st Century*, J Clark, ed. (Amsterdam: Harwood Academic Publishers), pp 77–101
- Georges M, Nielsen D, Mackinnon M, Mishra A, Okimoto R et al. (1995) Mapping quantitative trait loci controlling milk production by exploiting progeny testing. *Genetics* 139, 907–920
- Grobet L, Royo Martin LJ, Poncelet D, Pirottin D, Brouwers B et al. (1997) A deletion in the myostatin gene causes double-muscling in cattle. *Nat Genet* 17, 71–74
- Grobet L, Poncelet D, Royo Martin LJ, Brouwers B, Pirottin D et al. (1998) Molecular definition of an allelic series of mutations disrupting the myostatin function and causing double-muscling in cattle. *Mamm Genome* 9, 210–213
- Hetzel J, Brascamp P, Leveziel H, Lewin H, Teale A et al. (1993) The International Bovine Reference Panel (IBRP). Report of the ISAG working party.
- Hudson TJ, Stein LD, Gerety SS, Ma J, Castle AB et al. (1995) An STS-based map of the human genome. *Science* 270, 1945–1954, with supplementary data from the Whitehead Institute/MIT Center for Genome Research, Human Genetic Mapping Project, data release 11.9 (May 1997).
- Kambadur R, Sharma M, Smith TPL, Bass JJ (1997) Mutations in myostatin (GDF8) in double-muscling Belgian Blue Cattle. *Genome Res* 7, 910–916
- Kappes SM, Keele JW, Stone RT, McGraw RA, Sonstegard TS et al. (1997) A second-generation linkage map of the bovine genome. *Genome Res* 7, 235–249
- Lander E, Green P (1987) Construction of multilocus genetic linkage maps in humans. *Proc Natl Acad Sci USA* 84, 2363–2367
- Libert F, Lefort A, Okimoto R, Georges M (1993) Construction of a bovine genomic library of large yeast artificial chromosome clones. *Genomics* 18, 270–276
- Lunetta KL, Boehnke M, Lange K, Cox DR (1995) Experimental design and error detection for polyploid radiation hybrid mapping. *Genome Res* 5, 151–163
- Masabanda J, Ewald D, Buitkamp J, Potter K, Fries R (1998) Molecular markers for the bovine fibrillin 1 gene (FBN1) map to 10q26. *Anim Genet* in press
- McPherron AC, Lee SJ (1997) Double muscling in cattle due to mutations in the myostatin gene. *Proc Natl Acad Sci USA* 94, 12457–12461
- McPherron AC, Lawler AM, Lee SJ (1997) Regulation of skeletal muscle mass in mice by a new TGF $\beta$  superfamily member. *Nature* 387, 83–90
- Smith TPL, Lopez-Corrales NL, Kappes SM, Sonstegard TS (1997) Myostatin maps to the interval containing the bovine mh locus. *Mamm Genome* 8, 742–744
- Solinas Toldo S, Lengauer C, Fries R (1995a) Comparative genome map of human and cattle. *Genomics* 27, 489–496
- Solinas Toldo S, Mezzelani A, Hawkins GA, Bishop MD, Olsaker I et al. (1995b) Combined Q-banding and fluorescence in situ hybridisation for the identification of bovine chromosome 1 to 7. *Cytogenet Cell Genet* 69, 1–6
- Sonstegard TS, Lopez-Corrales NL, Kappes SM, Beattie CW, Smith TPL (1997) Comparative mapping of human chromosome 2 identifies segments of conserved synteny near the bovine mh locus. *Mamm Genome* 8, 751–755
- Takeda H, Yamakuchi H, Ihara N, Hara K, Watanabe T et al. (1998) Construction of a bovine yeast artificial chromosome (YAC) library. *Anim Genet* 23, 216–219
- Visscher PM, Van der Beek S, Haley CS (1998) Marker assisted selection. In *Animal Breeding: Technology for the 21st Century*, J Clark, ed. (Amsterdam: Harwood Academic Publishers), pp 119–136
- Walter MA, Spillet DJ, Thomas P, Weissenbach J, Goodfellow PN (1994) A method for constructing radiation hybrid maps of whole genomes. *Nat Genet* 7, 22–28
- Weir BS (1996) Genetic data analysis II. Methods for discrete population genetic data. (Sunderland, Mass.: Sinauer Associates, Inc.)
- Womack JE, Johnson JS, Owens EK, Rexroad CE 3rd, Schlapfer J et al. (1997) A whole-genome radiation hybrid panel for bovine gene mapping. *Mamm Genome* 8, 854–856